Modèles, primitives et méthodes de suivi pour la segmentation vasculaire. Application aux coronaires en imagerie tomodensitométrique 3D.

David Lesage

École Doctorale
d'Informatique,
TÈlÈcommunications
et Électronique de Paris

Thèse

présentée pour obtenir le grade de docteur

de l'Ecole Nationale Supérieure des Télécommunications

Spécialité : Signal et Images

David LESAGE

Modèles, Primitives et Schémas d'Extraction pour la Segmentation Vasculaire.

Application aux Coronaires en Imagerie Tomodensitométrique 3D.

Soutenue le 15 octobre 2009 devant le jury composé de

Laurent Cohen	**Rapporteurs**
Wiro Niessen	
Renaud Keriven	**Examinateurs**
Maciej Orkisz	
Isabelle Bloch	**Directrice de thèse**
Elsa Angelini	**Co-directrice de thèse**
Gareth Funka-Lea	**Co-encadrant de thèse**

David LESAGE

Models, Features and Extraction Schemes for Vascular Segmentation :

Application to the delineation of Coronary Arteries from 3D Computed Tomography Data.

Résumé

Les maladies vasculaires font partie des principaux problèmes de santé des pays développés. La segmentation des structures vasculaires dans les images médicales est une étape fondamentale pour l'aide au diagnostic, pour le traitement et la planification préopératoire, car elle simplifie considérablement les tâches de visualisation et de quantification. Malheureusement, une majorité de routines angiographiques cliniques repose encore sur des étapes manuelles. La quantité de données générées par les modalités d'imagerie 3D angiographiques telles que l'imagerie tomodensitométrique (Computed Tomography, CT) et l'imagerie par résonance magnétique (Magnetic Resonance, MR), rend la segmentation manuelle une tâche particulièrement fastidieuse et délicate. Dans ce contexte, les techniques de traitement d'images, automatiques ou semi-automatiques, ont pour but de faciliter et accélérer la tâche des experts médicaux en minimisant les interactions manuelles tout en reduisant la variabilité inter-opérateurs.

La segmentation des structures vasculaires en imagerie 3D est un problème difficile et spécifique à différents égards. Les vaisseaux sanguins forment des réseaux fins et complexes, et les acquisitions sont de plus en plus détaillées. Dans ce manuscript, nous présentons tout d'abord un état de l'art étendu des méthodes existantes en segmentation vasculaire 3D. Les avancées récentes sont discutées suivant trois axes d'étude : les modèles géométriques et d'apparence, les primitives extraites des images et les schémas d'extraction. La décomposition d'une méthode en composants individuels permet de plus facilement analyser points forts et limites, dans le but de proposer d'éventuelles améliorations.

Ces travaux se focalisent sur une problématique particulièrement complexe, la segmentation des artères coronaires en imagerie tomodensitométrique 3D (CT). Nous détaillons tout d'abord notre modèle géométrique axial, que nous évaluons dans l'image grâce à une primitive fondée sur le flux de gradient. Les principales caractéristiques de cette nouvelle primitive sont son fort pouvoir discriminant et son faible coût calculatoire.

Dans un deuxième temps, nous proposons un modèle bayésien récursif intégrant de la connaissance a priori sur les variations de notre modèle géométrique. Les différents termes de ce modèle Bayésien, a priori et vraisemblances, sont appris via une approche non-paramétrique, exploitant une base de segmentations manuelles de référence.

En exploitant les primitives et modèles précédents, nous proposons enfin deux schémas d'extraction différents. Notre premier schéma d'extraction reformule le problème de segmentation comme une procédure d'optimisation discrète sur graphe. Il met en oeuvre une approche fondée sur la propagation de chemins minimaux 4D (position spatiale 3D de l'axe du vaisseau plus estimation du rayon associé). En particulier, notre nouvel algorithme exploite une nouvelle métrique cumulative, que nous dérivons empiriquement de notre modèle bayésien. Notre deuxième schéma d'extraction repose sur une approche de suivi stochastique par méthode de Monte-Carlo séquentielle, également connue sous le nom de filtre particulaire. La distribution de probabilité a posteriori de notre modèle

bayésien est estimée séquentiellement grâce à l'évolution d'une population d'échantillons aléatoires. Pour ces deux schémas d'extraction, une attention particulière est donnée à la robustesse des résultats et à l'efficacité de la mise en oeuvre. Nous proposons notamment différentes stratégies et raffinements algorithmiques pour la fouille parcimonieuse de larges espaces de recherche. La robustesse des approches proposées est démontrée à travers une évaluation qualitative et quantitative sur une base large de données cliniques de grande taille et grande variété.

Summary

Vascular diseases are among the most important public health problems in developed countries. The segmentation of vascular structures is a fundamental step for diagnosis assistance, treatment and surgery planning, as it greatly eases visualization and quantification tasks. Unfortunately, most angiographic clinical routines still rely heavily on manual operations. Given the amount of data generated by modern 3D imaging modalities, *e.g.*, computed tomography angiography (CTA) and magnetic resonance angiography (MRA), manual segmentation has become a considerable burden. Automatic and semi-automatic image processing techniques aim at easing and speeding up reviewing tasks by medical professionals, reducing the amount of manual interaction and lowering inter-operator variability.

Vascular segmentation is a specific and challenging problem at different levels. Vessels are thin and complex structures embedded in increasingly large images. We first present an extensive bibliographical review of existing methods for 3D vascular segmentation. We discuss state-of-the-art developments with regards to three main axes of study : geometric and appearance models, image features and extraction schemes. Decomposing a method into individual components provides valuable insight on its eventual advantages and limitations, on which points it might be refined or extended towards better performance.

In this work, we focus on a particularly challenging problem, the delineation of coronary arteries from 3D cardiac Computed Tomography Angiographic (CTA) data. We first detail our medial-based geometric model, which we evaluate on the image thanks to a fast, oriented, gradient flux-based medialness feature.

We then formulate a recursive Bayesian model, embedding prior knowledge on the variations of the geometric model. We discuss how the different terms of our model, likelihood and priors, are learned in a non-parametric fashion from a ground-truth database of manually segmented datasets.

We propose two separate extraction schemes to perform the actual segmentation. Our first strategy derives from the reformulation of the problem as a discrete, graph-based optimization task. The extraction is performed thanks to minimal-path-like techniques, propagating paths on a 4D (position+radius) graph. We detail notably the derivation of an empirical cumulative cost metric from our seminal Bayesian formulation. Our second strategy employs a sequential, stochastic Monte-Carlo tracking approach, also known as particle filtering. It successively estimates the posterior distribution of our Bayesian model through a population of random samples. For both approaches, particular attention is given to robustness and computational efficiency. Different algorithmic refinements are proposed to explore large search spaces in a sparse and efficient manner. Qualitative and quantitative evaluation is presented on a varied database of clinical data.

Remerciements / Acknowledgments

This work was founded Siemens Corporate Research (Princeton NJ, USA), in collaboration with Telecom ParisTech (Paris, France).

I would like to thank warmly all the members of my committee who kindly accepted to review my thesis. Many thanks to Laurent Cohen and Wiro Niessen for their detailed and constructive comments and to Maciej Orkisz for presiding my committee. I am truly honored by their interest in my work. I was lucky enough to have three advisors during these years, Elsa Angelini, Isabelle Bloch and Gareth Funka-Lea. I can not thank them enough for their scientific insight, their suggestions and for their infinite patience. I am sure they must have hated me at times, yet they showed me only unconditional professional and human support. Gareth, thank you so much for your trust and guidance during these years in Princeton. You surely deserve one of these 'Best boss in the world' mugs. *Isabelle, tu as littéralement porté cette thèse sur tes épaules dans les moments difficiles. Ta compréhension, ta disponibilité et ta capacité à motiver même les doctorants les plus déprimés vont bien au-delà des limites humaines normales.*

I would like to thank all my colleagues and friends from SCR in Princeton. Working in such a cosmopolitan atmosphere was simply an incredible life and work experience. As an unnamed SCR intern said one day, 'you can never really forget about Princeton'. Thanks to this cute – somehow too cute – little town, I made friends from all around the world. Many thanks to all the amazing people I met during these years : Rama, Leo, Chris, Marie-Pierre, Tom, Chef, Timo, Mikael, Kamih, Mike, Major, the Herve, the Julien(s), Brad, Sean, Raj, Kinda, Jean-Marc, Yasmine, Alicia, Christophe, Pauline (no order of preference) and all the others I forgot. I would especially like to thank Matthias and Romain, my first two managers at SCR. I started this adventure a bit because of you.

Merci à toute l'équipe du département TSI de Télécom ParisTech, et tout particulièrement à Patricia pour sa gentillesse et sa bonne humeur légendaires. Je tiens évidemment à remercier la Medical Team et associés, parmi lesquels Jerem', Oliv' et Camille, qui ont partagé avec moi, au quotidien, les délices d'une fin de thèse. Merci aussi à Djo et Emi d'avoir supporté avec diplomatie mes incursions inopinées, une bonne dizaine de fois par jour. A ma décharge, votre porte était toujours ouverte ! Un grand bravo également pour le self-control de mes compagnons d'infortune, Nico (tu peux le faire !), Vincent et Giovanni notamment, qui ont eu la chance de partager la C01 avec moi. Merci encore à J.B., Carolina, Julien, Lazar et aux anciens, Gero, Pau, Antonio et Céline, pour leur bonne humeur et leur sympathie. Des remerciements particuliers vont aux notables de la Butte, parmi lesquels Ceyhun (le Barbare), Jamal (le Président) et, évidemment, Jérôme (le Roi). Nos discussions ont souvent été d'une importance que je n'ai réalisée que bien après. Que ce soit sur la vie, l'univers ou tout le reste (comme l'optimisation du modèle de Potts), nous

étions presque toujours au diapason.

Il convient également de remercier les amis qui ont souffert avec moi au cours de cette expérience, c'est-à-dire Bloup, Marion, Camille, Julien, Soizic, Nes, Francis et tous ceux que j'oublie. Milles excuses pour mes absences, mes comportements sociaux pas toujours adaptés et pour mon obsession compulsive pour tout ce qui ressemble de près ou de loin à une artère. Merci pour votre amitié sans faille et votre écoute, même si vous n'aviez pas toujours une idée très claire de ce dont j'étais en train de parler. Je tiens à vous rassurer, moi non plus. Je tiens évidemment à remercier chaleureusement ma famille, mes parents et mon frère, pour leur soutien indéfectible, pas seulement au cours de cette thèse. Enfin, mes derniers remerciements iront à celle qui a vécu au plus proche de l'épicentre. A celle qui n'avait aucune idée de ce qui l'attendait quand elle m'a rencontré et qui s'est révélée un ange de patience et de compréhension. A celle qui y a souvent cru plus que moi. Merci de tout mon coeur, Vanessa.

Contents

Chapitre 1

Présentation détaillée

1.1 Chapitre 2 : Introduction et contexte

De par son importance médicale et les problématiques complexes et spécifiques qu'elle pose, la segmentation vasculaire en imagerie médicale jouit d'une communauté scientifique particulièrement active. Dans le contexte du traitement d'images, les contributions de cette thèse se situent à différents niveaux, théoriques et pratiques. Nous abordons tout d'abord des problématiques de modélisation, autrement dit, l'injection de connaissances *a priori* dans le but d'améliorer la robustesse et la précision de la segmentation. Dans un deuxième temps, nous étudions la conception de primitives dédiées, c'est-à-dire l'exploitation d'informations image pertinentes et leur extraction pratique. Enfin, nous nous intéressons aux schémas d'extraction, combinant modèles, primitives et algorithmes d'optimisation pour parvenir au résultat de segmentation. Tout en suivant ces axes méthodologiques, ces travaux de thèse proposent des contributions pratiques à une application de premier ordre, la segmentation des artères coronaires en imagerie tomodensitométrique 3D (CT en anglais).

Cette thèse a été réalisée en partenariat entre Télécom ParisTech, Paris, France, et Siemens Corporate Research, Princeton NJ, Etats-Unis, sous la co-direction d'Isabelle Bloch et Elsa Angelini (Télécom ParisTech) et Gareth Funka-Lea (Siemens Corporate Research).

1.1.1 Contexte médical

Les artères coronaires, irriguant le muscle cardiaque (figure 2.1), jouent un rôle majeur dans l'appareil circulatoire et sont à l'origine de graves et fréquentes pathologies (Rosamond et al., 2008; World Health Organization, 2008) telles que la *Maladie Coronarienne* (Coronary Heart Disease, *CHD*).

L'imagerie tomodensitométrique 3D (CT) est une alternative moderne, de haute résolution et non-invasive aux routines angiographiques par rayons X pour le diagnostic, la planification chirurgicale et le suivi des pathologies coronariennes. Parmi les inconvénients de la technologie CT, on peut mentionner les doses de radiations potentiellement élevées et la taille des données 3D engendrées, rendant leur examen manuel fastidieux et délicat. Différentes techniques de visualisation permettent de faciliter la tâche du clinicien (voir section 2.1). Malheureusement, ces techniques nécessitent généralement une segmentation préalable, ou tout du moins une délinéation initiale des artères d'intérêt. La conception d'outils informatiques permettant de faciliter au maximum cette étape est donc un enjeu clinique de premier ordre.

1.1.2 Problématiques

Ces travaux explorent les problématiques liées à la conception d'une méthode de segmentation vasculaire, s'appuyant sur le cas d'étude des artères coronaires en imagerie CT. Dans ce contexte, nous considérons la délinéation initiale des vaisseaux d'intérêt comme une étape primordiale, qui conditionne de façon critique la robustesse et la précision du résultat final.

En partant de l'acquisition originale, le but de ces travaux est l'extraction robuste d'une représentation compacte et de haut niveau du réseau vasculaire d'intérêt. Une telle représentation facilite grandement l'interprétation et la manipulation des données originelles et peut servir de base à de nombreux raffinements ultérieurs (amélioration ultérieure de la précision de la segmentation, quantification de structures pathologiques et génération de visualisations dédiées).

En termes de traitement d'images, la délinéation des artères coronaires en imagerie CT est une application particulièrement complexe :

- les jeux de données CT sont volumineux (de l'ordre de 512^3 voxels) ;
- inversement, les artères d'intérêt sont généralement fines, variant typiquement entre 1 et 10 voxels de rayon dans de telles acquisitions ;
- les données sont affectées par le bruit et les artefacts d'acquisition, l'hétérogénéité de l'agent de contraste et les effets de volume partiel ;
- la géométrie des arbres coronaires est particulièrement variable et complexe ;
- différentes pathologies (calcifications, stents, anévrismes et sténoses) peuvent altérer de façon significative la géométrie et l'apparence locales des artères ;
- diverses structures hyper-intenses (chambres cardiaques, veines) sont présentes dans le voisinage immédiat des artères coronaires.

Le partenariat industriel avec Siemens Corporate Research a motivé le développement d'outils pratiques, utilisables dans des scénarios cliniques réalistes. Une attention particulière est portée sur la robustesse et la vitesse d'exécution des techniques présentées. Nous détaillons différentes contributions algorithmiques et d'implantation et validons nos travaux sur une base de données cliniques de grande taille et de grande variété.

1.1.3 Structure du document

Le chapitre 3 propose un état de l'art étendu des travaux existants en segmentation vasculaire 3D. Cette étude bibliographique est organisée suivant trois axes méthodologiques : modèles, primitives et schémas d'extraction.

Les chapitres 4 à 8 reprennent ces axes d'étude pour la présentation des contributions originales de cette thèse. Le chapitre 4 introduit un modèle axial pour la représentation géométrique des structures vasculaires et discute différentes propriétés photométriques (modèle d'apparence) des artères coronaires en imagerie CT.

Dans le chapitre 5, nous introduisons une nouvelle primitive, `MFlux` , que nous utilisons pour l'évaluation du modèle géométrique (position axiale, rayon et orientation) sur les données images. Cette primitive dérive d'une mesure du flux de gradient au travers de sections circulaires (`Flux`). Ses propriétés théoriques et pratiques sont discutées et comparées à une sélection de techniques existantes.

Le chapitre 6 présente notre modèle bayésien pour les artères coronaires en imagerie CT. Par sa formulation *a posteriori*, ce modèle combine de l'information géométrique *a priori* et l'information extraite de l'image. Nous expliquons comment nous construisons

les différentes composantes de ce modèle par apprentissage, à partir d'une base de segmentations de référence.

En exploitant les primitives et modèles mentionnés précédemment, nous proposons enfin deux schémas d'extraction différents. Notre premier schéma d'extraction, présenté dans le chapitre 7, met en œuvre une procédure d'optimisation discrète sur graphe propageant des chemins minimaux 4D (position axiale 3D plus estimation du rayon associé). L'originalité de cette approche réside notamment dans l'utilisation d'une nouvelle métrique cumulative, que nous dérivons empiriquement de notre modèle bayésien. Notre deuxième schéma d'extraction, détaillé dans le chapitre 8, repose sur une approche de suivi stochastique multi-hypothèses. Nous utilisons une variante de filtre particulaire (méthode de Monte-Carlo séquentielle) pour l'estimation des états successifs de notre modèle bayésien évalué sur les données images. Pour ces deux schémas d'extraction, une attention particulière est donnée à la robustesse des résultats et à l'efficacité de la mise en œuvre. Nous proposons notamment différentes stratégies et raffinements algorithmiques pour la fouille parcimonieuse de larges espaces de recherche. La robustesse des approches proposées est démontrée à travers une évaluation qualitative et quantitative sur une base données cliniques de grande taille.

1.2 Chapitre 3 : Etat de l'art

Le chapitre 3 est une version adaptée et mise à jour de l'article (Lesage et al., 2009a), publié dans le journal Medical Image Analysis. Il propose un état de l'art étendu des travaux existants en segmentation vasculaire, se focalisant plus particulièrement sur les modalités d'imagerie 3D (CT et MR) et sur la segmentation de la lumière vasculaire.

Etant donné le très vaste champ d'applications vasculaires, notre revue bibliographique adopte un point de vue haut niveau. Nous présentons les grandes tendances méthodologiques du domaine et discutons les innovations récentes suivant trois axes d'étude :

- les *modèles* d'apparence et de géométrie (section 3.1) ;
- les *primitives* extraites des images (section 3.2) ;
- les *schémas d'extraction* menant à la segmentation finale (section 3.3).

Au travers de ces axes d'étude, nous tentons d'éclairer certaines questions clés rencontrées lors de la conception d'une méthode de segmentation vasculaire :

- quelle connaissance *a priori* est exprimée ?
- comment cette connaissance est-elle exploitée, évaluée sur les données images ? Quelle est la fiabilité de ces mesures ?
- comment l'information extraite des images est-elle contrôlée, corrigée, optimisée vis-à-vis des modèles ?

Une méthode de segmentation vasculaire, dans sa globalité, correspond souvent à une combinaison plus ou moins complexe de modèles, primitives et schémas d'extraction (voir table 3.1). La décomposition d'une méthode en composants individuels permet une meilleure compréhension de ses fondements théoriques et de ses propriétés pratiques, et donc de ses points forts, éventuelles limites et potentielles améliorations.

1.2.1 Modèles d'apparence et de géométrie

La section 3.1 discute les principaux modèles d'apparence et de géométrie vasculaire utilisés dans la littérature. Nous distinguons :

- les modèles d'*apparence* (section 3.1.1), exploitant de l'information *a priori* sur les propriétés photométriques des vaisseaux d'intérêt, leur voisinage et la modalité d'acquisition.
- les modèles *géométriques* (section 3.1.2), exploitant la forme spécifique des vaisseaux sanguins, notamment leur élongation. Parmi les modèles géométriques, une attention particulière est portée sur les modèles *médiaux* et les *cylindres généralisés*. Ceux-ci architecturent leur représentation autour d'une *ligne centrale*, une courbe médiale encodant de façon simple et efficace le centre de la lumière vasculaire.
- les modèles dits *hybrides*, proposant une modélisation jointe de l'apparence et de la géométrie. On retrouve dans cette catégorie les modèles d'*apparence radiale*, reposant sur une géométrie tubulaire pour décrire les variations photométriques le long de rayons de sections circulaires.

Nous mentionnons finalement différents travaux portant sur la modélisation des bifurcations, des pathologies vasculaires et sur l'utilisation de modèles spécifiques à certaines applications.

Le choix d'un modèle vasculaire adéquat dépend évidemment de nombreux facteurs. Les propriétés photométriques (et donc le modèle d'apparence) sont très largement dépendantes de la modalité d'acquisition visée. L'*a priori* géométrique exploitable varie sensiblement suivant l'application anatomique (taille, régularité et complexité du réseau vasculaire d'intérêt) mais également suivant les buts applicatifs. Par exemple, les modèles les plus simples ne satisfont généralement pas les degrés de précision requis pour une quantification pathologique fine. Il est évidemment possible de combiner différents modèles pour améliorer à la fois la robustesse et la précision d'une méthode de segmentation. Une approche classique est alors l'utilisation de modèles de plus en plus fins, utilisés successivement à différents stades de raffinement de la segmentation.

1.2.2 Primitives vasculaires

Dans la section 3.2, nous présentons différentes catégories de *primitives vasculaires* de la littérature, c'est-à-dire les détecteurs et filtres spécifiques utilisés pour évaluer un modèle vasculaire sur des données images.

Nous distinguons notamment :

- les primitives *isotropes*, n'exploitant pas l'élongation présumée d'un vaisseau sanguin ; les techniques de cette catégorie reposent généralement sur une hypothèse d'hyper- ou hypo-intensité locale du vaisseau d'intérêt ;
- les primitives exploitant la *géométrie 3D locale*, incluant notamment les méthodes dérivatives telles que l'étude de la Hessienne multi-échelle, particulièrement populaire, les méthodes intégratives fondées sur l'étude des moments d'inertie, et l'optimisation locale d'un modèle hybride ;
- les primitives *2D* évaluant les *sections orthogonales* des vaisseaux d'intérêt ; ces méthodes reposent généralement sur l'évaluation de la position médiale et d'une détection plus ou moins grossière de la frontière de la lumière vasculaire.

Nous étudions également différents efforts de recherche concernant la détection de bifurcations et de pathologies telles que les sténoses, les anévrismes et les calcifications.

Les primitives existantes diffèrent en termes de précision, robustesse et vitesse de calcul. Une autre considération importante est la quantité d'information qu'une primitive peut extraire de l'image. Par exemple, les détecteurs multi-échelle fondés sur l'étude de la Hessienne évaluent à la fois l'orientation et l'échelle pour une position donnée, mais

peuvent être sensibles au bruit et aux anomalies locales. Les primitives 2D de section orthogonale permettent en général des estimations robustes de l'échelle (rayon) locale, mais requièrent une estimation préalable de l'orientation locale du vaisseau. Les paramètres d'entrée peuvent être vus comme définissant l'espace d'optimisation du schéma d'extraction. L'utilisation de primitives spécialisées, présentant de nombreux paramètres, engendre généralement des processus d'optimisation plus complexes. Néanmoins, l'incorporation d'information de haut niveau, contrôlée par le schéma d'extraction, peut permettre de compenser les problèmes de fiabilité des informations de bas niveau extraites de l'image. Le contrôle de cet équilibre haut niveau / bas niveau est une des problématiques clé au cœur de la conception d'une méthode de segmentation vasculaire.

1.2.3 Schémas d'extraction

La section 3.3 est consacrée à l'état de l'art des schémas d'extraction, c'est-à-dire les algorithmes mis en œuvre pour réaliser la segmentation en pratique. Nous discutons tout d'abord différentes techniques de pré-traitement et de pré-segmentation permettant d'augmenter la qualité des acquisitions (filtrage de rehaussement anisotrope, par exemple), simplifier leur contenu (*e.g.* par sous-échantillonnage) ou obtenir une localisation approximative des vaisseaux d'intérêt (masques et régions d'intérêts). De tels résultats préliminaires peuvent alors servir d'initialisation aux schémas d'extraction proprement dits. Nous distinguons :

- les algorithmes de *croissance de région* (section 3.3.2) ; cette catégorie comprend la croissance de région classique, fondée sur des critères d'inclusion voxel par voxel, mais aussi les algorithmes de *propagation par vague ou par front* où la cohérence spatiale du processus est contrôlée ; de tels algorithmes procèdent de manière itérative et progressive, permettant notamment l'inclusion de critères adaptatifs améliorant la robustesse et la précision du résultat final ;
- les méthodes de *contours actifs* (section 3.3.3) ; cette catégorie inclut un certain nombre de méthodes paramétriques (*snakes*) spécifiquement adaptés aux structures fines telles que les vaisseaux sanguins, ainsi que des approches implicites telles que des méthodes par *ensembles de niveaux* (*level-sets*) dédiés à la segmentation vasculaire ;
- les approches fondées sur la *ligne centrale* (section 3.3.4) ; ces algorithmes visent avant tout à extraire la structure médiale des vaisseaux cibles ; nous distinguons les méthodes par suivi direct, qui segmentent itérativement le vaisseau en suivant pas-à-pas sa ligne centrale, et les processus d'optimisation tels que les méthodes par *chemins minimaux* ; nous discutons notamment de récents développements sur lesquels s'appuient certains des travaux originaux de cette thèse ;
- de récentes avancées dans le cadre de méthodes stochastiques multi-hypothèses (section 3.3.5), incluant des travaux sur les *processus de Markov pointés* et les *filtres particulaires* (méthode de Monte-Carlo séquentielle) ; ces dernières méthodes sont particulièrement attractives de par leur capacité à combiner modèles et informations extraites de l'image dans un cadre théorique flexible et puissant.

Nous présentons enfin différentes techniques de post-traitement (section 3.3.6) permettant de raffiner les résultats de segmentation (amélioration de la robustesse aux faux positifs et faux négatifs et amélioration de la précision), mais également de changer de représentation finale (obtention d'une ligne centrale à partir d'une segmentation du contour et inversement).

Parmi les critères de choix d'un schéma d'extraction, on peut naturellement mentionner la vitesse d'exécution requise ainsi que le degré d'interaction et le type de résultat désiré. Une fois encore, il peut être particulièrement bénéfique de combiner plusieurs schémas d'extraction. La figure 3.11 récapitule la structure de cette section et illustre certaines des combinaisons possibles entre les différentes techniques présentées.

1.3 Chapitre 4 : Modèles de géométrie et d'apparence

Le chapitre 4 présente les modèles utilisés dans nos travaux pour incorporer de l'information *a priori* sur la géométrie et l'apparence spécifique des structures vasculaires en imagerie CT 3D.

1.3.1 Modèle Géométrique : représentation médiane discrète

Notre modèle géométrique 3D repose sur une représentation médiane discrète. Il est directement inspiré des travaux sur les représentations médiales de formes générales, visant à encoder une forme au travers d'une collection structurée d'axes médiaux, formant le "squelette" de la forme. Depuis les travaux de (Blum, 1967) sur la transformation en axe médian (*Medial Axis Transform*, MAT), les propriétés théoriques et pratiques des représentations médianes ont été étudiées dans de nombreux travaux (Brady and Asada, 1984; Naf et al., 1996; Pizer et al., 1998; Fridman, 2004; Pizer et al., 1998, 2003). Ces modèles jouissent notamment d'une grande compacité. Une forme 2D ou 3D est simplement représentée grâce à ces axes médians (courbes 1D ou surfaces 2D) et une valeur de rayon associée (rayon du cercle ou de la sphère inscrit(e) correspondant à une position d'un axe).

Dans la section 3.1, nous avons exposé brièvement l'intérêt de tels modèles pour décrire la géométrie 3D de structures vasculaires. Si dans le cas de formes générales, un modèle médian correspond à une surface 2D, il est en général raisonnable d'approcher une structure vasculaire 3D à l'aide d'une courbe 1D correspondant à sa *ligne centrale* (voir figure 4.2). De tels modèles géométriques sont particulièrement séduisants de par leur simplicité et sont de ce fait populaires dans le domaine de la segmentation vasculaire. Les modèles fondés sur la ligne centrale exploitent directement l'élongation spécifique des structures vasculaires en imposant une topologie 1D. Les modèles proposés dans la littérature diffèrent notamment suivant la définition de l'axe central (courbes linéaires par morceaux, B-splines ou splines cardinales, par exemple) et la définition de l'information volumique associée. Une approche directement inspirée des représentations médiales classiques est d'associer à la ligne centrale une valeur de rayon correspondant à une sphère inscrite ou à une section vasculaire supposée circulaire. Les modèles les plus avancés proposent une définition plus fine des sections des vaisseaux au travers d'ellipses ou de courbes paramétriques, voire une définition de la paroi vasculaire par une surface 2D contrainte vis-à-vis de la ligne centrale.

Des modèles vasculaires médiaux existants, notre modèle hérite sa simplicité et sa compacité. L'une de ses principales caractéristiques est sa paramétrisation discrète, qui simplifie encore son expression tout en assurant un fort pouvoir expressif (voir figure 4.3). La ligne centrale est discrétisée sous forme d'une séquence de points centraux $\{p_t\}_{t=0,\dots,L}$. A chaque point central est associé une valeur de rayon r_t et une direction tangentielle d_t, qui permettent de définir une section circulaire centrée en p_t. Notre modèle peut donc être entièrement décrit par des séquences de triplets $x_{0:L} = \{(p_t, r_t, d_t)\}_{t=0,\dots,L}$. Un tel triplet x_t correspond à un élément atomique ou *état* de notre modèle discret.

Une des originalités de notre modèle est la définition des directions tangentielles $\{d_t\}$ au travers de *points de contrôle* $\{c_t\}$

$$d_t = \frac{c_t - c_{t-1}}{\|c_t - c_{t-1}\|}$$

et de contraindre le positionnement des points de contrôle vis-à-vis des points centraux $\{c_t\}$:

$$p_t = \frac{c_t + c_{t-1}}{2}$$

Cette contrainte couple indirectement les points centraux et les directions tangentielles d'une manière similaire aux splines cardinales. Ce couplage permet une définition stable des directions tangentielles même en cas de forte courbure. Il induit également une réduction de la dimensionalité du modèle via sa reparamétrisation, puisque les états $x_t = (p_t, r_t, d_t)$ de notre modèle peuvent également être décrits suivant $x_t = (c_{t-1}, c_t, r_t)$, c'est-à-dire à partir de la séquence de points de contrôle et les rayons associés.

Malgré sa paramétrisation discrète, la ligne centrale de notre modèle peut aisément être interpolée, en utilisant par exemple un modèle de spline hermitienne cubique. Notre modèle peut également incorporer des branchements, comme illustré dans la figure 4.5.

La simplicité de notre modèle repose notamment sur l'utilisation de sections circulaires induisant un unique paramètre de rayon. Cette restriction est jugée comme empiriquement raisonnable pour les structures vasculaires relativement fines telles que les artères coronaires, comme illustré dans la figure 4.4.

Un paramètre clé de notre modèle est son pas de discrétisation, c'est-à-dire l'espacement des points de contrôles successifs $s_t = \|c_{t-1} - c_t\|$. Ce paramètre affecte directement le pouvoir expressif du modèle. Dans ces travaux appliqués aux artères coronaires en imagerie CT, nous utilisons un pas fixe correspondant à la résolution intra-coupe typique de nos données ($s - s_t = 0,3\text{mm}, \forall t$). Dans la section 4.1.3, nous démontrons que ce pas de discrétisation fin permet une capture précise des variations typiques des artères coronaires. Par une simple procédure d'optimisation itérative, nous avons évalué les erreurs angulaires induites par notre modèle sur une base de segmentations manuelles (annexe B). Celles-ci se révèlent très limitées (moyenne de $0,078 \pm 0,011$ radians et maximum de $0,23$ radians), confirmant le bon pouvoir expressif de notre modèle. L'interpolation, par spline cubique hernitienne, de la ligne centrale, de la direction tangentielle et du rayon est également d'une précision satisfaisante (voir table 4.2). Enfin, les modèles géométriques optimisés sur notre base de segmentations manuelles sont directement réutilisés pour l'apprentissage des différents termes de notre modèle Bayésien, décrit dans le chapitre 6.

1.3.2 Modèle d'apparence : echelle Hounsfield en imagerie CT

Dans une acquisition tomodensitométrique 3D (CT), l'intensité de chaque voxel correspond à une valeur d'atténuation (radio-densité) intégrée sur l'unité volumétrique correspondante. Ces valeurs d'intensité, aussi appelées *nombres CT*, sont exprimés suivant une échelle d'unité arbitraire dénommée unité Hounsfield (H.U.). L'origine (zéro) de cette échelle correspond traditionnellement à la valeur d'atténuation de l'eau. Pour une échelle Hounsfield ainsi calibrée, des tissus de coefficients d'atténuation différents occupent donc des intervalles d'intensité distincts et prévisibles (figure 4.7). Notre application se focalisant sur les artères coronaires, nous pouvons concentrer notre analyse sur un intervalle d'intensité restreint, correspondant aux tissus sanguins rehaussés par agent de contraste.

Nous appliquons aux acquisitions le pré-traitement suivant :

$$\tilde{I}(x) = \begin{cases} T_{low} & \text{si } I(x) < T_{low} \\ T_{high} & \text{si } I(x) > T_{high} \\ I(x) & \text{sinon} \end{cases} \tag{1.1}$$

où $I(x)$ est la valeur CT originale d'un voxel x et $\tilde{I}(x)$ sa valeur transformée. Les seuils T_{low} et T_{high} restreignent la dynamique des valeurs dans l'intervalle $[T_{low}, T_{high}]$. Nous choisissons de façon conservative $T_{low} = -24H.U.$ et $T_{high} = 576H.U.$. Le seuil bas permet d'exclure les poumons, mais inclut au moins partiellement les tissus mous tels que le myocarde. Le seuil haut sépare le sang rehaussé de tissus de haute densité tels que les calcifications. Les tissus dont les intensités résident dans l'intervalle $[T_{low}, T_{high}]$ correspondent potentiellement aux artères coronaires (sang rehaussé) et ne sont donc pas altérés.

La figure 4.8 illustre l'effet de ce pré-traitement. Premièrement, l'interface poumon-cœur est largement atténuée, ce qui réduit les risques de faux positifs détectés par nos primitives image (chapitre 5). Deuxièmement, l'intensité des calcifications est réduite à celle du seuil haut T_{high}. En pratique, l'une des conséquences de cette opération est l'inclusion potentielle des calcifications dans nos résultats de segmentation. Les calcifications ne font pas partie de la lumière vasculaire et sont souvent explicitement exclues dans la littérature sur la segmentation des coronaires (Friman et al., 2008b; Metz et al., 2008a). Nous motivons notre choix par l'observation de nombreux cas de sclérose aigüs, où la lumière vasculaire visible est quasiment inexistante. Dans de telles situations, les calcifications représentent des indices structurels primordiaux. Dans nos expériences, leur inclusion plutôt que leur exclusion permet un gain de robustesse considérable pour nos schémas d'extraction (chapitres 7 et 8). Des traitements supplémentaires seraient néanmoins nécessaires pour exclure les zones calcifiées *a posteriori* et ainsi obtenir une véritable segmentation de la lumière vasculaire.

1.3.3 Modèle d'apparence radiale : profil de barre convoluée

Les modèles d'apparence radiale décrivent les variations d'intensité le long de profils radiaux de sections vasculaires (voir la section 3.1.3 de notre étude bibliographique). Bien que nous nous n'exploitions pas explicitement de tels modèles dans nos méthodes, nous utilisons le modèle par barre convoluée (Krissian et al., 2000; Wörz and Rohr, 2004, 2007, 2008) pour l'étude de nos primitives images au travers de tests artificiels (section 5.4).

Le profil de barre convoluée correspond à un plateau convolué par un noyau gaussien (figure 4.9). Le plateau modélise le vaisseau idéal (état stationnaire de l'agent de contraste) tandis que la convolution gaussienne modélise la fonction d'atténuation spécifique de l'imageur. Dans la section 4.3, nous illustrons notamment que ce modèle permet de refléter fidèlement l'apparence des artères coronaires en imagerie CT. Contrairement à des modèles plus simples tels que le profil gaussien, il permet de rendre compte de l'apparence variable des artères en fonction de leur diamètre : un profil quasi gaussien pour les plus fines, l'apparition d'un plateau pour les plus larges (voir figure 4.9).

L'optimisation du modèle de barre convoluée sur notre base de segmentations manuelles nous a notamment permis de déterminer les paramètres adéquats pour la générations des tests artificiels de la section 5.4.

1.4 Chapitre 5 : Primitives fondées sur le flux de gradient : `Flux` et `MFlux`

Le chapitre 5 est une version mise à jour et étendue de travaux publiés dans IEEE International Symposium on Biomedical Imaging (ISBI 2009) (Lesage et al., 2009c). Dans ce chapitre, nous présentons la primitive utilisée pour évaluer notre modèle géométrique dans les images. Celle-ci, dénommée `MFlux` , est dérivée d'une mesure locale de flux de gradient (primitive `Flux`), modifiée pour obtenir un plus haut pouvoir discriminant. Nous proposons une implantation efficace de ces primitives et évaluons leurs propriétés théoriques et pratiques au travers d'une série d'expériences sur des données réelles et synthétiques.

Nos primitives `Flux` et `MFlux` sont destinées à un usage de détection, d'évaluation des différents paramètres (position axiale, rayon et orientation) de notre modèle. En ce sens, elles sont donc conçues pour être intégrées à des méthodes de segmentation telles que des approches articulées autour de l'extraction de la ligne centrale (voir chapitres 7 et 8). Il est important de noter que ce but applicatif diffère sensiblement d'applications telles que le rehaussement par filtrage, notamment en termes d'exigences de sensibilité et de précision.

1.4.1 Description de `Flux` et `MFlux`

La norme (amplitude) des gradients d'une image est une mesure classiquement utilisée pour la détection de contours. Au-delà de cette norme, les méthodes fondées sur le flux de gradient exploitent l'orientation des vecteurs de gradient en mesurant le flux au travers de l'interface de l'objet segmenté. Pour les objets hyper-intenses, le flux entrant est maximisé lorsque le contour (en 2D) ou la surface (en 3D) extrait(e) est alignée avec le champ de gradient de l'image (voir figure 5.1).

Le flux de gradient fut notamment exploité pour la segmentation vasculaire dans les travaux de (Vasilevskiy and Siddiqi, 2002; Descoteaux et al., 2004; Kimmel and Bruckstein, 2003; Holtzman-Gazit et al., 2006) au sein de méthodes fondées sur les ensembles de niveaux (level-sets). Ces travaux ont notamment démontré que les mesures de flux de gradient sont bien adaptées à l'extraction de structures fines, alongées et faiblement contrastées. Dans le contexte de surfaces générales telles que celles optimisées par les méthodes par ensembles de niveaux, certains travaux tels que (Vasilevskiy and Siddiqi, 2002) ont proposé d'approcher la mesure globale du flux via des mesures locales, au travers de sphères multi-échelles. Dans le cas particulier de notre modèle géométrique, nous pouvons nous dégager de cette approximation en considérant les sections circulaires de notre modèle comme autant d'éléments de surfaces locaux. Cette observation mène à la formulation de notre première primitive, `Flux` .

1.4.2 `Flux`

Les paramètres locaux (p, r, d) (position axiale, rayon et direction tangentielle) de notre modèle géométrique permettent de définir un contour de section $C(p, r, d)$. La primitive `Flux` mesure la contribution locale du flux de gradient à travers un contour de test $C(p, r_{\texttt{Flux}}, d)$ dont le rayon $r_{\texttt{Flux}}$ est défini en fonction du rayon r du modèle (voir section 5.4.3). Après avoir discrétisé le contour circulaire $C(p, r_{\texttt{Flux}}, d)$ en N points $\{x_i\}_{i=1,\ldots,N}$, nous obtenons la formulation de `Flux` :

$$\texttt{Flux}(p, r_{\texttt{Flux}}, d) = \frac{1}{N} \sum_{i=1}^{N} \langle \nabla I(x_i), u_i \rangle \tag{1.2}$$

où $u_i = \frac{p-x_i}{\|p-x_i\|}$ est la direction radiale unitaire intérieure (voir figure 5.2). Ces directions radiales, dérivées directement des paramètres de notre modèle, sont utilisées comme approximations des normales à la surface vasculaire.

1.4.3 `MFlux`

La primitive `MFlux` dérive d'une modification de `Flux` introduite pour atténuer les réponses erronées aux contours isolés, un inconvénient commun à de telles mesures linéaires (voir figure 5.3). En particulier, la primitive `Flux` a tendance à répondre fortement le long des structures hyper-intenses larges telles que les chambres cardiaques dans nos données CT. De tels faux positifs sont caractérisés par des contributions asymétriques des gradients à la mesure de flux (voir figure 5.3). Inspirés par l'idée initiale de (Koller et al., 1995), nous proposons une combinaison non-linéaire des gradients diamétralement opposés :

$$\texttt{MFlux}(p, r_{\texttt{MFlux}}, d) = \frac{2}{N} \sum_{i=1}^{\frac{N}{2}} \min(\langle \nabla I(x_i), u_i \rangle, \langle \nabla I(x_i^\pi), u_i^\pi \rangle)$$

où $x_i^\pi = x_{\frac{N}{2}+i}$ pour un nombre N pair de points de contour. Notre primitive `MFlux` retient donc uniquement la contribution minimale par paire de points de contour.

1.4.4 Relation avec des travaux existants

Parmi les travaux fondés sur des mesures de flux de gradient, nous avons déjà mentionné des mesures de flux au travers de sphères multi-échelle. Une implantation particulièrement efficace d'une telle primitive est notamment proposée dans (Law and Chung, 2009). En comparaison, nos primitives `Flux` et `MFlux` exploitent localement la direction tangentielle d donnée par notre modèle géométrique. Suivant des idées similaires, la primitive *optimally oriented flux* (OOF) de (Law and Chung, 2008) se fonde sur une mesure locale de l'orientation maximisant le flux de gradient. L'implantation fréquentielle proposée par (Law and Chung, 2008) est bien adaptée au filtrage (rehaussement) de volumes de données entiers étant donné un ensemble restreint de rayons de test pré-sélectionnés. En comparaison, nos primitives `Flux` et `MFlux` sont conçues pour tester de façon parcimonieuse des paramètres donnés (position axiale, rayon et orientation) et sont donc bien adaptés aux schémas d'extraction de type suivi.

Nos primitives `Flux` et `MFlux` utilisent un patron de test (section circulaire) similaire à d'autres primitives de la littérature, fondées sur les idées introduites par (Koller et al., 1995). Les travaux (Koller et al., 1995) utilisent des dérivées gaussiennes orientées suivant les directions radiales pour mesurer l'information contour le long du patron de test. Des idées similaires peuvent notamment être trouvées dans les travaux de (Xu and Pycock, 1999; Poli and Valli, 1996; Fridman, 2004). En comparaison, nos primitives fondées sur le flux de gradient projettent les vecteurs gradient sur les directions radiales. Un schéma numérique similaire peut notamment être trouvé dans (Krissian et al., 2000; Wink et al., 2000a; Pock et al., 2005). Il est particulièrement intéressant de noter les forts liens entre un schéma par projection des vecteurs gradient et un schéma par dérivées orientées. Dans un cadre gaussien continu, ils peuvent même s'avérer équivalents d'un point de vue théorique. Les travaux existants diffèrent notamment de par l'utilisation de schémas de régularisation variés. Certains travaux (Koller et al., 1995; Krissian et al., 2000; Wink et al., 2000a; Fridman, 2004; Pock et al., 2005) proposent d'adapter la force de la régularisation (gaussienne) en fonction de l'échelle de l'objet d'intérêt (rayon), suivant les principes

théoriques des espaces linéaires gaussiens multi-échelles (Lindeberg, 1994). Le choix de la force de régularisation résulte d'un compromis entre la robustesse de la détection et la précision de la localisation (Fridman, 2004). Pour l'implantation de nos primitives, nous avons opté pour une régularisation gaussienne pour sa simplicité et son efficacité calculatoire. D'autres approches, telles qu'un lissage anisotrope, pourraient être explorées. Nous choisissons également d'appliquer une régularisation fixe (indépendante du rayon) et relativement faible. Pour notre application à des données CT cardiaques de haute résolution et ayant un relativement bon rapport signal sur bruit, une telle approche permet d'obtenir un bon compromis entre robustesse et précision (Gulsun and Tek, 2008b). De plus, l'utilisation d'une régularisation spatialement limitée permet de réduire l'influence néfaste que pourraient avoir les structures hyper-intenses voisines sur les vaisseaux d'intérêt (Law and Chung, 2008). Finalement, le fait d'utiliser une régularisation fixe a d'importants avantages calculatoires. Notre approche nécessite un unique champ de gradient régularisé quel que soit le rayon testé, et ce champ peut être pré-calculé.

Nos primitives `Flux` et `MFlux` peuvent être considérées comme des primitives de type médial (*medialness features*), c'est-à-dire des primitives conçues en priorité pour la détection d'axes médians (ligne centrale du vaisseau). Suivant la classification de telles primitives proposée par (Pizer et al., 1998), les primitives que nous proposons peuvent être qualifiées de primitives *décentrées* (*offset*), c'est-à-dire exploitant l'information de *contour* le long du patron de test circulaire.

1.4.5 Détails d'implantation

La section 5.3 détaille notre implantation de `Flux` et `MFlux` . De par leur formulation discrète particulièrement simple, l'implantation de `Flux` et `MFlux` est directe et aisée. Cependant, une implantation naïve peut être grandement optimisée.

Pour notre application aux coronaires en imagerie CT, nous utilisons $N = 8$ points de contours. Ce choix est un compromis entre robustesse et rapidité de calcul. Pour notre schéma d'extraction par optimisation discrète sur graphe (chapitre 7), nous utilisons un ensemble discret, prédéfini de directions tangentielles, ce qui nous permet de pré-calculer certaines quantités telles que les directions radiales unitaires.

Le champ de vecteurs gradient est pré-calculé sur une grille 3D isotrope. Le volume de données original est tout d'abord rééchantillonné sur une grille isotrope par interpolation tri-linéaire. La résolution de cette grille est fixée à la résolution intra-coupe initiale des données (typiquement de l'ordre de $0,3mm$). Une transformation des intensités (voir section 4.2) est ensuite appliquée pour focaliser l'analyse sur l'intervalle d'intensité d'intérêt qui englobe notamment les tissus sanguins rehaussés par agent de contraste. Enfin, le champ de gradient est conjointement calculé et régularisé grâce à des convolutions par dérivées de gaussiennes. Nous exploitons l'implantation récursive rapide de (Deriche, 1992). L'écart-type utilisé est fixé à $\sigma_g = 0,3mm$ et permet une régularisation limitée ne dégradant que faiblement les artères les plus fines. Pour obtenir les vecteurs gradient à n'importe quel point de contour x_i requis par `MFlux` , nous utilisons une interpolation tri-linéaire du champ discret pré-calculé. Pour les rayons de tests relativement grands ($r_{\texttt{MFlux}}$ supérieur à 1,5mm), une interpolation au plus proche voisin peut être substituée pour une rapidité accrue sans dégradation notable de la précision.

Finalement, nous proposons différentes optimisation d'encodage et de stockage des données volumiques. Les valeurs de gradient sont quantifiées sur 8 bits pour réduire l'occupation mémoire au minimum. Nous proposons également une allocation mémoire par

bloc volumique (voir figure 5.4) favorisant une utilisation cohérente du cache processeur. Notre implantation de `Flux` et `MFlux` exploite de surcroît les instructions vectorisées `Intel SSE3`, permettant d'atteindre des performances de l'ordre de 6 millions (interpolation au plus proche voisin des vecteurs gradient) et 0,7 million (interpolation tri-linéaire) d'évaluations par seconde sur un processeur à 2,16 GHz. Ce niveau de performance élevé joue un rôle primordial dans l'efficacité finale de nos schémas d'extraction, qui requièrent typiquement des milliers voire des millions d'évaluations de réponse de `MFlux` pour segmenter un vaisseau d'intérêt.

1.4.6 Expériences

La section 5.4 illustre les propriétés théoriques et pratiques des primitives `Flux` et `MFlux` au travers d'une série d'expériences sur des données synthétiques et des données cliniques réelles. Nos primitives sont notamment comparées à une sélection de primitives de la litérature :

- la populaire primitive `Vesselness` de (Frangi et al., 1998), fondée sur l'analyse multi-échelle de la matrice Hessienne ;
- la primitive `Ribbon` proposée par (Florin et al., 2005), fondée sur une mesure de contraste dans des sections circulaires (voir figure 5.5) ;
- la primitive `Core` de (Fridman, 2004) fondée sur un patron de test circulaire similaire à `Flux` et `MFlux` , mais reposant sur des dérivées gaussiennes orientées et d'écart-type variable (voir figure 5.6).

Les données synthétiques sont générées suivant le modèle radial de barre convoluée (voir section 4.3), suivant des paramètres (contraste, rayons...) correspondant à notre application aux coronaires en imagerie CT. Les réponses des différentes primitives sont normalisées afin de faciliter leur comparaison.

Nous étudions notamment :

- la relation entre le paramètre d'échelle (rayon ou échelle de test optimal) et le rayon théorique du vaisseau ;
- la précision de l'estimation de la position de la ligne centrale du vaisseau ;
- la précision de l'estimation du rayon du vaisseau ;
- la sensibilité à l'orientation du vaisseau ;
- le comportement des différentes primitives pour des changements joints :
  - d'échelle et de position de test ;
  - d'échelle et d'orientation de test ;
  - d'orientation et de position de test ;
- l'influence du bruit d'image sur la réponse des différentes primitives ;
- la réponse en cas de branchements vasculaires, en présence de structures voisines hyper-intenses et en cas de pathologies telles que des sténoses ou des anévrismes ;
- l'influence du contraste.

Nous mettons notamment en évidence les propriétés particulièrement discriminantes de `Flux` et `MFlux` vis-à-vis du décentrement de la position centrale testée (point p). Comme illustré dans la figure 5.8, les réponses de nos primitives décroissent rapidement en fonction du décentrement. Comparativement aux autres primitives testées, leur pouvoir discriminant augmente avec le rayon du vaisseau d'intérêt. La primitive `MFlux` , grâce à son comportement non-linéaire pénalisant les contributions asymétriques au flux de gradient, accentue encore cet avantage. Il est à noter cependant qu'une primitive telle que `Vesselness` fut conçue initialement pour le rehaussement des vaisseaux par filtrage. Son

comportement moins discriminant vis-à-vis du décentrement en fait d'elle une primitive mieux adaptée à cet usage que `Flux` ou `MFlux` .

Nous illustrons également le bon comportement de `Flux` et `MFlux` vis-à-vis de l'estimation du rayon (figure 5.9). Grâce à l'utilisation d'une régularisation fixe et indépendante du rayon de test, la réponse de `Flux` et `MFlux` décroît plus rapidement que les autres primitives pour des rayons de test dépassant le rayon réel du vaisseau.

Les expériences portant sur les variations jointes de l'échelle et de la position de test illustrent le problème de faux positifs en présence de contours isolés (figure 5.11). Celui-ci apparaît de façon particulièrement marquée pour de faibles échelles de test et des positions excentrées pour les primitives `Flux` et `Core` et pour de grandes échelles pour les primitives `Vesselness` et `Core` . La composante non-linéaire introduite dans la formulation de `MFlux` réduit sensiblement ce problème. Le fort pouvoir discriminant de notre primitive est démontré de nouveau dans la figure 5.19, illustrant par des données synthétiques les problèmes potentiels causés par une large structure hyper-intense (par exemple, une chambre cardiaque) adjacente à un vaisseau d'intérêt. Des exemples de situations analogues sont donnés en figure 5.23 pour des volumes CT cardiaques cliniques.

En contrepartie de son fort pouvoir discriminant, `MFlux` apparaît comme l'une des primitives les plus sensibles au bruit (figure 5.16). Cette observation est cependant modérée par les niveaux de bruit relativement faibles rencontrés dans nos données CT cardiaques.

Les figures 5.18 et 5.20 illustrent, sur des cas synthétiques, le comportement satisfaisant de nos primitives en cas de bifurcations et d'anomalies anatomiques (sténoses et anévrismes). Des exemples supplémentaires de la réponse aux bifurcations sont donnés dans les figures 5.23, 5.24 et 5.25 sur des données cliniques. Comparativement à la primitive `Vesselness` notamment, notre primitive `MFlux` jouit d'une meilleure robustesse aux branchements et d'une meilleure réponse aux fines branches faiblement contrastées. Elle démontre également une meilleure localisation du pic de réponse au niveau de la ligne centrale du vaisseau.

Ces expériences démontrent que les primitives que nous proposons, `Flux` et surtout `MFlux` , se comparent favorablement à différents travaux existants en termes de pouvoir discriminant, de précision et de robustesse. Combinées à un faible coût calculatoire, ces bonnes propriétés théoriques et pratiques font de `MFlux` une primitive bien adaptée pour des méthodes de segmentation fondées sur l'extraction de la ligne centrale du vaisseau.

1.5 Chapitre 5 : Modèle bayésien pour les artères coronaires en imagerie tomodensitométrique 3D

Le chapitre 6 introduit notre modèle bayésien récursif, qui intègre de l'information a priori sur les variations de notre modèle géométrique (chapitre 4) et interprète les réponses de notre primitive `MFlux` (chapitre 5) en termes de distributions de vraisemblance.

1.5.1 Formulation récursive

Notre modèle géométrique définit un segment vasculaire comme une suite d'états discrets $x_t = (p_t, r_t, d_t)$ avec p_t points de ligne centrale, r_t les valeurs de rayon et d_t les directions tangentielles associées. Nous considérons désormais le modèle $x_{0:L}$ en tant que variable aléatoire. Une chaîne $x^i_{0:L} = \{(p^i_t, r^i_t, d^i_t)\}^L_{t=0}$ particulière correspond donc à une réalisation de cette variable aléatoire.

La formulation par maximum a posteriori (MAP) correspond à l'extraction de la réalisation de plus haute probabilité étant donné les données :

$$x^*_{0:L} = \arg\max_{\mathcal{X}_L} p(x_{0:L}|z_{1:L}) \tag{1.3}$$

où $z_{1:L} = \{z_t\}_{t=1}^{L}$ sont des observations tirées des données et disponibles à chaque pas de temps t.

Nos buts applicatifs divergent du problème MAP classique sur plusieurs points :

– la longueur L du vaisseau d'intérêt est a priori inconnue ;

– la structure vasculaire d'intérêt peut présenter plusieurs vaisseaux (arbre vasculaire).

L'espace de recherche explose généralement de façon exponentielle avec la longueur L de la chaîne. En pratique, il n'est donc possible d'explorer qu'une fraction très limitée de cet espace. Puisque nous sommes intéressés par l'extraction d'arbres vasculaires complets, notre analyse ne porte pas seulement sur l'unique réalisation de maximum a posteriori. Nous utilisons une formulation a posteriori pour comparer et sélectionner différentes réalisations de notre modèle.

La distribution a posteriori est donnée par le théorème de Bayes :

$$p(x_{0:L}|z_{1:L}) = \frac{p(x_{0:L})p(z_{1:L}|x_{0:L})}{p(z_{1:L})} \tag{1.4}$$

où $p(x_{0:L})$ est la distribution a priori et $p(z_{1:L}|x_{0:L})$ est la vraisemblance jointe des observations. Le terme d'a priori intègre de l'information sur les variations du modèle. Le terme de vraisemblance peut être considéré comme un terme d'attache aux données. La distribution marginale des observations $p(z_{1:L})$ est indépendante de la réalisation spécifique du modèle considérée, mais dépend de la longueur L du modèle.

Nos schémas d'extraction (chapitres 7 et 8) sont optimisés de façon itérative. Il est donc intéressant d'exprimer la distribution a posteriori sous une forme récursive Sous l'hypothèse d'un a priori markovien du premier ordre, l'a priori devient :

$$p(x_{0:L}) = p(x_0)\prod_{t=1}^{L} p(x_t|x_{t-1}) \tag{1.5}$$

Les observations $\{z_t\}$ sont considérées comme indépendantes étant donné le modèle :

$$p(z_{1:L}|x_{0:L}) = \prod_{t=1}^{L} p(z_t|x_t) \tag{1.6}$$

Ces hypothèses sont classiquement adoptées dans le cadre de problèmes d'estimation bayésienne récursive (Doucet et al., 2000). En combinant ces équations, on obtient une forme récursive de la distribution a posteriori (Doucet et al., 2000) :

$$\begin{aligned}
p(x_{0:t}|z_{1:t}) &= \frac{p(x_{0:t})p(z_{1:t}|x_{0:t})}{p(z_{1:t})} \\
&= p(x_0)\prod_{k=1}^{t}\frac{p(x_k|x_{k-1})p(z_k|x_k)}{p(z_{1:t})} \\
&= \frac{p(x_t|x_{t-1})p(z_t|x_t)}{p(z_t|z_{1:t-1})}p(x_0)\prod_{k=1}^{t-1}\frac{p(x_k|x_{k-1})p(z_k|x_k)}{p(z_{1:t-1})} \\
&= \frac{p(x_t|x_{t-1})p(z_t|x_t)}{p(z_t|z_{1:t-1})}p(x_{0:t-1}|z_{1:t-1})
\end{aligned} \tag{1.7}$$

La distribution a posteriori est mise à jour de $t-1$ à t au travers de l'a priori de transition $p(x_t|x_{t-1})$ et de la vraisemblance marginale $p(z_t|x_t)$. Ces deux termes permettent de définir notre processus markovien.

1.5.2 Apprentissage non-paramétrique

Pour définir les différents termes de notre modèle bayésien, nous exploitons une base de 61 jeux de données CT cardiaques segmentés manuellement (voir annexe B). Cette base nous fournit un large ensemble de réalisations de différentes quantités telles que les rayons, directions tangentielles et réponses de `MFlux` le long d'artères coronaires. Grâce à cette précieuse source d'information statistique, nous proposons d'apprendre les différents termes de notre modèle bayésien au travers d'un schéma d'estimation de densité non-paramétrique.

Etant donné deux variables aléatoires Z_1 et Z_2, notre schéma d'estimation peut être décrit comme suit :

1. la densité $\hat{f}(Z_1, Z_2)$ est estimée à partir d'échantillons joints $\{(Z_1^i, Z_2^i)\}$;
2. la densité conditionnelle $p(Z_2|Z_1 = z_1)$ est extraite.

Ce schéma général est illustré dans la figure 6.1. Les densités jointes de type $\hat{f}(Z_1, Z_2)$ capturent les variations jointes de Z_1 et Z_2. Elles sont apprises sur notre base d'échantillons grâce à une méthode non-paramétrique d'estimation par noyaux (Silverman, 1986). De ces densités jointes, nous pouvons dans un deuxième temps extraire les différentes densités conditionnelles nécessaires à la définition de notre modèle. Ce schéma est appliqué à l'apprentissage des distributions de vraisemblances comme des a priori composant notre modèle. Grâce à cette méthode, nous sommes en mesure d'estimer de riches densités continues de notre base d'échantillons discrets. Ce processus d'estimation est dit non-paramétrique car il n'impose pas de modèle de distribution de probabilité analytique.

1.5.3 Vraisemblance marginale $p(z_t|x_t)$

La section 6.3 détaille la décomposition de notre modèle d'observation pour les artères coronaires en imagerie CT, c'est-à-dire la définition de la distribution marginale de vraisemblance $p(z_t|x_t)$ de notre modèle bayésien.

A chaque pas de temps t, nous considérons l'ensemble $z_t = \{y_t^j\}_{j=1}^{N_t}$ des observations disponibles. Une observation y_t^i correspond à la réponse de notre primitive `MFlux` (chapitre 5) associée à une réalisation x_t^i de notre modèle. En d'autres termes, la primitive `MFlux` est évaluée pour une position p_t^i, un rayon r_t^i et une direction tangentielle d_t^i.

Le terme $p(z_t|x_t)$ correspond à la vraisemblance jointe de toutes les observations étant donné l'état du modèle au temps t. Tout d'abord, nous faisons l'hypothèse que les observations sont conditionnellement indépendantes étant donné le modèle :

$$p(z_t|x_t) = \prod_{j=1}^{N_t} p(y_t^j|x_t) \tag{1.8}$$

Cette hypothèse permet la simplification du modèle, le rendant optimisable en pratique.

Une des particularités de notre modèle d'observation est de considérer deux types de distributions de vraisemblance, l'une pour les observations issues des vaisseaux d'intérêts $(p_v(y_t|r_t = r_t^i),)$ et l'autre pour les observations issues du fond $(p_{bg}(y_t))$. L'exploitation explicite de l'information extraite du fond a d'importants avantages en termes de sélectivité

et de robustesse (Geman and Jedynak, 1996; Konishi et al., 2003; Yuille and Coughlan, 1997, 2000). Nous exploitons le fait que les observations sont associées à des réalisations d'états du modèle :

$$p(y_t^j|x_t = x_t^i) = \begin{cases} p_v(y_t^i|x_t = x_t^i) & \text{si } j = i \\ p_{bg}(y_t^j) & \text{sinon} \end{cases} \tag{1.9}$$

Etant donnée une réalisation d'état x_t^i, une observation y_t^j est considérée comme venant d'une distribution de vraisemblance de vaisseau avec la probabilité $p_v(y_t^i|x_t = x_t^i)$ si cette observation est effectivement associée à l'état x_t^i $(j = i)$. Sinon $(j \neq i)$, l'observation est considérée comme appartenant au fond avec la probabilité $p_{bg}(y_t^j)$. Les distributions $p_v(y_t|x_t)$ dépendent du modèle. Elles varient notamment en fonction du rayon r_t considéré.

L'expression de la vraisemblance marginale peut être décomposée suivant :

$$p(z_t|x_t = x_t^i) = p_v(y_t^i|x_t = x_t^i) \prod_{j \neq i} p_{bg}(y_t^j) \tag{1.10}$$

Une interprétation de cette équation est que les différentes réalisations du modèle considérées sont mises en compétition pour expliquer toutes les observations à disposition (Geman and Jedynak, 1996). Cette expression peut-être re-factorisée sous la forme :

$$p(z_t|x_t = x_t^i) = \frac{p_v(y_t^i|x_t = x_t^i)}{p_{bg}(y_t^i)} \prod_{y_t^j \in z_t, \forall j} p_{bg}(y_t^j) \tag{1.11}$$

Le premier terme, sous forme de ratio, encode la compétition entre les vraisemblances du vaisseau et du fond pour une observation y_t^i associée à l'état x_t^i. Le second terme est un produit sur toutes les observations, indépendant de la réalisation du modèle. En d'autres termes, cette expression permet de focaliser notre analyse uniquement sur les observations associées à la réalisation du modèle considérée.

Les distributions de vraisemblance de réponse dans les vaisseaux $p_v(y_t|x_t)$ sont apprises sur notre base de segmentations manuelles. Comme illustré dans la figure 6.2, nous estimons tout d'abord la densité jointe $\hat{f}_v(y_t, r_t)$ des réponses de `MFlux` en fonction du rayon. Nous pouvons ensuite extraire de cette densité des distributions conditionnelles de type $p_v(y_t|r_t = r_t^i)$ en considérant une réalisation particulière de rayon r_t^i (voir figure 6.4). Notre schéma d'estimation non-paramétrique permet de capturer les variations subtiles de réponses de `MFlux` en fonction du rayon.

La distribution de vraisemblance du fond $p_{bg}(y_t)$ est construite d'une façon légèrement différente, suivant les principes proposés par (Konishi et al., 2003). Pour chaque nouveau jeu de données, nous échantillonnons de façon aléatoire l'espace de paramètres (p_t, r_t, d_t) et évaluons les réponses de `MFlux` correspondantes. L'utilisation d'un grand nombre d'échantillons (de l'ordre de 10^5 évaluations de `MFlux`) permet l'obtention de statistiques stables. L'idée principale est que, pour un tel échantillonnage, la probabilité de tirer des paramètres correspondant à de véritables vaisseaux est très faible (voir (Konishi et al., 2003) pour une analyse formelle). Nous pouvons donc considérer ces échantillons comme provenant du fond. Nous obtenons tout d'abord des densités jointes $\hat{f}_{bg}(y_t, r_t)$ de réponse en fonction du rayon. L'hypothèse est faite que ces réponses sont approximativement indépendantes du rayon, comme illustré dans la figure 6.3. Nous obtenons donc une distribution $p_{bg}(y_t)$ dont la plus importante caractéristique est d'être spécifique au jeu de données considéré, variant notamment en fonction du rapport signal sur bruit.

Notre modèle d'observation encode la compétition entre les deux types de distribution de vraisemblance pour l'explication d'une observation. Comme illustré dans la figure 6.4, ces distributions sont généralement bien séparées, confirmant le fort pouvoir discriminant de notre primitive `MFlux` . Les zones de recouvrement encodent l'ambiguïté potentielle dans l'interprétation des observations. Il est à noter que notre modèle d'observation exploite implicitement le fait que notre modalité de choix, l'imagerie tomodensitométrique (CT), est supposée calibrée. Notre méthodologie devrait donc être modifiée pour être appliquée à des modalités non calibrées telles que l'imagerie par résonance magnétique.

1.5.4 A priori de transition $p(x_t|x_{t-1})$

Le dernier terme de notre modèle bayésien restant à définir est l'a priori de transition markovien $p(x_t|x_{t-1})$, encodant de l'information sur les variations potentielles de notre modèle géométrique. Nous décomposons ce terme en fonction des variations de rayon et de direction tangentielle :

$$p(x_t|x_{t-1}) = p(p_t, r_t, d_t|p_{t-1}, r_{t-1}, d_{t-1}) = p(r_t|r_{t-1})p(d_t|d_{t-1}, r_{t-1}) \tag{1.12}$$

La section 6.4 discute les hypothèses permettant cette décomposition.

Les distributions d'a priori de transition de rayon $p(r_t|r_{t-1})$ sont également apprises de façon non-paramétrique sur notre base de segmentations manuelles, en considérant comme échantillons des réalisations de rayons successifs (r_t^i, r_{t-1}^i) (voir figures 6.5 et 6.6). On peut noter que les zones de forte densité de probabilité se situent le long d'un axe $r_t = r_{t-1}$, confirmant l'intuition que les variations de rayon sont relativement lentes le long des artères coronaires. Notre processus d'estimation non-paramétrique permet de capturer certaines variations subtiles, telles qu'une variance légèrement plus élevée pour les rayons larges, et le caractère asymétrique des distributions $p(r_t|r_{t-1} = r_{t-1}^i)$. Ces distributions illustrent l'observation que les rayons des coronaires tendent à décroître entre les ostia et leurs parties distales.

Le terme d'a priori directionnel $p(d_t|d_{t-1}, r_{t-1})$ est défini en fonction des changements angulaires :

$$p(d_t|d_{t-1}, r_{t-1}) = p(\alpha_t|r_{t-1}) \tag{1.13}$$

avec $\alpha_t = \arccos d_t \cdot d_{t-1} = \arccos \frac{(c_{t-1}-c_{t-2})\cdot(c_t-c_{t-1})}{\|c_{t-1}-c_{t-2}\|\|c_t-c_{t-1}\|}$ (voir figure 6.7). Les distributions de type $p(\alpha_t|r_{t-1} = r_{t-1}^i)$ sont une nouvelle fois extraites d'une densité jointe $\hat{f}(\alpha_t, r_{t-1})$ apprise de façon non-paramétrique sur notre base de vérité terrain (voir figure 6.8). On peut notamment observer qu'intuitivement, les artères coronaires semblent conserver des trajectoires relativement stables (maxima de probabilité pour des angles faibles). Ces modes de densité apparaissent cependant pour des angles non-nuls, tendant à prouver que les trajectoires des coronaires ne sont généralement pas tout à fait rectiligne. On observe également une plus forte variabilité angulaire pour les plus faibles rayons. Une nouvelle fois, la capture de telles subtilités est rendue possible par notre processus d'apprentissage non-paramétrique combiné à la richesse statistique disponible dans notre base de segmentations manuelles.

1.6 Chapitre 7 : Extraction par optimisation de chemins minimaux

Le chapitre 7 est une version étendue de travaux publiés dans Medical Image Computing and Computer Assisted Intervention (MICCAI 2009) (Lesage et al., 2009b). Nous présentons notre premier schéma d'extraction, fondé sur une optimisation discrète sur graphe. Sur un graphe 4D (positions spatiales de la ligne centrale et valeur de rayon), notre stratégie propage et optimise des réalisations $x^i_{1:t}$ de notre modèle discret. Notre méthode emploie une nouvelle métrique cumulative, que nous déduisons empiriquement de notre modèle bayésien. Un accent particulier est porté sur l'efficacité algorithmique. Nous introduisons notamment différentes stratégies pour l'exploration efficace de si larges espaces de recherche et détaillons notre implantation. L'un des avantages de notre cadre algorithmique est sa grande souplesse d'utilisation, depuis un usage semi-interactif reposant sur une optimisation contrainte entre deux points jusqu'à l'extraction automatique d'arbres vasculaires complets. L'excellente robustesse de notre approche est démontrée par une étude qualitative et quantitative sur une base de données cliniques de grande taille.

1.6.1 Les chemins minimaux pour la segmentation vasculaire

Comme discuté précédemment dans notre revue bibliographique (section 3.3.4), l'optimisation de chemins minimaux est une approche particulièrement populaire pour la segmentation vasculaire fondée sur l'extraction de lignes centrales. L'idée centrale de telles approches est de formuler l'extraction de la ligne centrale comme une optimisation de chemins étant donnée une métrique cumulative. Le résultat de cette optimisation correspond au chemin de coût cumulé minimum (Cohen and Kimmel, 1997).

Les méthodes existantes peuvent différer dans leurs stratégies de discrétisation de l'espace de recherche (généralement, une grille cartésienne couvrant l'espace 3D), mais, surtout, dans la définition de la métrique cumulative employée. Pour les applications à la segmentation vasculaire, la métrique est généralement conçue pour pénaliser les chemins déviant des lignes centrales d'intérêt. Il est donc classique de définir cette métrique comme réciproque de primitives dédiées aux structures vasculaires, afin d'obtenir des coûts élevés pour des réponses de primitives faibles. Une des caractéristiques clés d'une métrique cumulative est la norme utilisée pour son intégration numérique. Les méthodes discrètes sur graphes telles que les travaux de (Gulsun and Tek, 2008b; Wink et al., 2004) correspondent généralement à une optimisation de norme L_1 et utilisent classiquement des variantes de l'algorithme de Dijkstra (Dijkstra, 1959). L'utilisation de la norme L_1 a généralement l'avantage d'une grande efficacité calculatoire, mais peut introduire des erreurs de métrification liées à sa nature discrète (Deschamps, 2001). L'algorithme de *fast-marching* (Adalsteinsson and Sethian, 1995; Sethian, 1996a; Tsitsiklis, 1995) est un schéma d'optimisation cohérent avec la formulation continue du problème de chemin minimal (Deschamps, 2001; Cohen and Kimmel, 1997). Il approche le véritable coût cumulatif de norme euclidienne L_2 et est de ce fait moins dépendant de la discrétisation employée. En contrepartie, son coût calculatoire est généralement plus élevé que celui des méthodes de type Dijkstra. Les méthodes fondées sur l'algorithme de *fast-marching* sont particulièrement populaires dans le domaine de la segmentation vasculaire (Avants and Williams, 2000; Deschamps and Cohen, 2001; Law and Heng, 2000; Lin, 2003; Wink et al., 2000b, 2001; Young et al., 2001). Enfin, on peut mentionner les approches de type "connexité floue" (Choyke et al., 2001; Udupa et al., 2002; Yim et al., 2000, 2003) qui correspondent à une optimisation

discrète de la norme L_∞.

Les problèmes d'optimisation de chemins minimaux ont été largement étudiés dans la littérature et jouissent de plusieurs propriétés particulièrement attractives :

1. il existe des algorithmes pour résoudre ce type de problèmes de façon globalement optimale. Pour la segmentation vasculaire, de tels algorithmes ont notamment démontrer leur grande robustesse, même en cas de sténoses ou d'occlusions (Lin, 2003) ou plus généralement de déconnexions partielles (Cohen and Deschamps, 2001) ;
2. les algorithmes existants sont généralement très efficaces d'un point de vue calculatoire. La plupart des méthodes de chemins minimaux, reposant sur des variantes de l'algorithme de Dijkstra ou du *fast-marching*, partagent une conception similaire, fondée sur une propagation ordonnée de sous-chemins optimaux ;
3. ce cadre algorithmique propose une grande souplesse d'utilisation, depuis un usage interactif (optimisation entre deux points) jusqu'à un usage entièrement automatique (exploration exhaustive de l'espace de recherche). Les algorithmes de chemins minimaux peuvent en général facilement accommoder des raffinements (Lin, 2003).

Parmi les problèmes classiquement rencontrés dans ce type d'approches, on peut tout d'abord mentionner les choix de la discrétisation de l'espace de recherche et du système de voisinage (définition de la structure du graphe sur lequel l'optimisation est réalisée). Ces choix peuvent sensiblement influencer la qualité du résultat final (voir section 7.1.3 et (Deschamps, 2001)). Un deuxième type de problème courant est l'observation d'une tendance au "décentrage" et au "raccourci" (voir figure 7.3). Différentes approches heuristiques ont été proposées pour réduire ce type de problèmes (Deschamps and Cohen, 2000; Deschamps, 2001; Deschamps et al., 1999). De tels phénomènes sont liés à la nature cumulative de l'optimisation, qui favorise par nature des chemins de faible longueur. Une approche générale pour limiter ces effets négatifs est de porter une attention particulière à la conception de la métrique cumulative. La robustesse d'une méthode est notamment étroitement liée au pouvoir discriminant des primitives employées : plus ces primitives seront discriminantes, plus la méthode sera robuste aux phénomène de décentrage et de raccourcis.

Certains travaux récents ont montré que l'espace de recherche n'était pas nécessairement limité à une grille 3D spatiale. Par l'ajout de dimensions supplémentaires telles que le rayon vasculaire (Wink et al., 2004; Li and Yezzi, 2007; Benmansour et al., 2009; Pechaud et al., 2009) ou l'orientation (Kimmel and Sethian, 2001; **?**; Pechaud et al., 2009), ces méthodes permettent l'optimisation de modèles plus complexes dans des cadres algorithmiques similaires. Les méthodes multi-échelles (ligne centrale + rayon) ont notamment démontré une plus grande robustesse aux problèmes de décentrage et autres ambiguïtés relatives à l'échelle d'analyse (voir figures 7.4 et 7.5). Une contrepartie est un coût calculatoire plus élevé (l'espace de recherche explose avec l'ajout de nouvelles dimensions) ainsi que l'apparition de différents problèmes de conception de la métrique cumulative (pondération entre termes spatiaux et termes d'échelle, par exemple).

Une dernière tendance récente est le développement de méthodes anisotropes, favorisant la propagation de chemins dans certaines directions (Lin, 2003; Benmansour et al., 2009; Benmansour and Cohen, 2009). Les directions locales favorisées sont typiquement dérivées d'estimations locales de la direction du vaisseau. De telles approches démontrent une robustesse accrue aux problèmes de raccourcis (Lin, 2003). Les travaux de (Benmansour et al., 2009; Benmansour and Cohen, 2009) proposent notamment une approche anisotrope et multi-échelle particulièrement élégante appliquée à la segmentation vasculaire 3D. Son

application reste malheureusement relativement coûteuse. Alternativement, les travaux de (Wink et al., 2004; Gulsun and Tek, 2008b) proposent d'exploiter, sur un graphe discret, la direction des arcs pour l'évaluation des coûts cumulatifs (voir figure 7.6). Nous réutilisons cette idée dans notre approche.

1.6.2 Du modèle bayésien à une métrique cumulative

La section 7.2 décrit la conception de notre nouvelle métrique cumulative, que nous déduisons empiriquement de notre modèle bayésien.

Notre modèle géométrique est exprimé comme une série discrète $x_{1:L}$ d'états $x_t = (p_t, d_t, r_t)$, de position centrale p_t, direction tangentielle d_t et rayon r_t. Notre modèle bayésien a posteriori est formulé de manière récursive et intègre un ensemble d'observations $z_t = \{y_t^i\}$ liées aux réalisations d'états x_t^i. Il est possible de lier notre formulation bayésienne à une optimisation cumulative en notant que maximiser la probabilité a posteriori de notre modèle est équivalent à minimiser l'opposé de son logarithme (Rousson et al., 2006). Le problème de maximum a posteriori entre les états x_0 et x_f peut être reformulé comme la recherche de la réalisation $x^*_{1:L^*}$ telle que :

$$x^*_{1:L^*} = \arg\min_{\mathcal{X}_{x_0,x_f}} \left(-\log(p(x_{0:L}|z_{1:L})) \right) \tag{1.14}$$

Dans notre cadre applicatif, la longueur optimale L^* est inconnue. L'ensemble $\mathcal{X}_{x_0,x_f}$ des réalisations possible comprend des candidats de longueurs différentes. Le développement de $-\log(p(x^i_{0:t}|z^i_{1:t})$ pour notre modèle est donné par :

$$\begin{aligned} -\log(p(x^i_{0:t}|z_{1:t})) &= -\log(p(x^i_{0:t-1}|z_{1:t-1})) - \log(p(x^i_t|x^i_{t-1})) - \log(p(z_t|x^i_t)) \\ &\quad + \log(p(z_t|z_{1:t-1})) \\ &= -\log(p(x^i_{0:t-1}|z_{1:t-1}) - \log(p(r^i_t|r^i_{t-1})) - \log(p(d^i_t|d^i_{t-1}, r^i_{t-1})) \\ &\quad - \log\left(\frac{p_v(y^i_t|r^i_t)}{p_{bg}(y^i_t)}\right) - \sum_{y^j_t \in z_t, \forall j} \log(p_{bg}(y^j_t)) + \log(p(z_t|z_{1:t-1})) \end{aligned} \tag{1.15}$$

Cette expression est additive et strictement croissante avec la longueur t de la chaîne, la rendant potentiellement utilisable dans un cadre d'optimisation de chemins minimaux. Différents problèmes se posent cependant :

- certain termes, tels que $\log(p(z_t|z_{1:t-1}))$ et la somme des vraisemblances $\sum_{y^j_t \in z_t, \forall j} \log(p_{bg}(y^j_t))$, ne peuvent pas être calculés de manière analytique ;
- l'a priori directionnel $\log(p(d^i_t|d^i_{t-1}, r^i_{t-1}))$ pose des problèmes de discrétisation (voir section 7.3).

La métrique cumulative que nous proposons est une simplification de l'équation 1.15. Nous définissons le coût $E(x^i_{0:t})$ d'une réalisation $x^i_{0:t}$:

$$E(x^i_{0:t}) = E(x^i_{0:t-1}) + C(x^i_{t-1}, x^i_t) \tag{1.16}$$

avec $C(x^i_{t-1}, x^i_t)$ le coût de mise à jour entre l'état x^i_{t-1} et l'état x^i_t. Celui-ci est défini par :

$$C(x^i_{t-1}, x^i_t) = -\log(p(r^i_t|r^i_{t-1})) - \log\left(\frac{p_v(y^i_t|r^i_t)}{p_{bg}(y^i_t)}\right) + M \tag{1.17}$$

avec M une constante positive discutée ultérieurement. Comparativement à l'équation 1.15, notre coût de mise à jour retient uniquement les termes dépendants de la réalisation

$x^i_{0:t}$ considérée. L'a priori directionnel est absent, pour des raisons discutées dans la section 7.3.

La constante M est fixée de façon à ce que le coût de mise à jour reste positif. Cette contrainte est nécessaire pour obtenir une métrique croissante, indispensable à l'utilisation d'un cadre algorithme efficace tel que la programmation dynamique. Nous fixons :

$$M = \max_{r_{t-1},r_t} \log(p(r_t|r_{t-1})) + \max_{r_t,y_t} \log(p_v(y_t|r_t)) - p_{min} \tag{1.18}$$

en imposant une borne inférieure $p_{min} = 10^{-6}$ sur toutes les probabilités. La section 7.2.3 discute de manière approfondie l'interprétation de la constante M, notamment en termes de normalisation vis-à-vis de la longueur du chemin considéré. En choisissant M comme la valeur minimale assurant un coût de mise à jour positif, nous pénalisons au minimum la longueur du chemin, ce qui permet notamment une meilleure robustesse aux problèmes de décentrage et de raccourcis.

Le coût d'une réalisation / chemin est calculé cumulativement par l'addition de nouveaux états x^i_t à $x^i_{0:t-1}$. Intuitivement, le coût $C(x^i_{t-1}, x^i_t)$ dépend seulement de l'état précédent x^i_{t-1} et comporte deux parties :

- un terme d'a priori sur les variations de rayon : $-\log(p(r^i_t|r^i_{t-1}))$;
- un terme d'attache aux données, correspondant au logarithme d'un ratio de vraisemblances : $-\log\left(\frac{p_v(y^i_t|r^i_t)}{p_{bg}(y^i_t)}\right)$.

Notre métrique cumulative mène naturellement à l'optimisation de chemins multi-échelles, similairement à (Li and Yezzi, 2007; Benmansour et al., 2009). Là où Benmansour et al. (2009) introduisent des paramètres empiriques pour contrôler les variations de rayons vis-à-vis des déplacements spatiaux, notre formulation exploite un a priori issu de notre modèle bayésien. On peut noter que tous les termes de notre métrique, a priori comme vraisemblances, sont appris automatiquement sur notre base de segmentations manuelles.

Enfin, la nature discrète de notre modèle géométrique et notre modèle bayésien ont naturellement mené à l'expression d'une métrique additive sous forme de somme discrète sur les états successifs. Nous proposons d'optimiser le coût cumulatif de réalisations de notre modèle, considérés en tant que chemins propagés sur un graphe. Cela correspond à une optimisation de chemin minimal de norme L_1.

1.6.3 Structure de graphe et de voisinage

La section 7.3 détaille la discrétisation de l'espace de recherche engendré par notre modèle géométrique. L'une des particularités de notre modèle est la définition de points de contrôle couplant points centraux et directions tangentielles (voir chapitre 4). Les états peuvent être définis de façon équivalente suivant $x_t = (p_t, d_t, r_t) \equiv (c_{t-1}, c_t, r_t)$, ce qui a permis de réduire la paramétrisation intrinsèque de notre modèle à 4 dimensions, des points de contrôle 3D et des valeurs de rayon. Pour discrétiser cet espace, nous proposons donc de définir un graphe 4D composé de deux parties :

- une grille spatiale 3D discrétisant les positions possibles de points de contrôle ;
- un ensemble discret de valeurs de rayon considérées.

La grille spatiale est définie par une grille de type "Body-Centered Cubic" (BCC) isotrope muni d'un voisinage de 14-connexité (voir figure 7.7). En comparaison à des grilles plus classiques telles que des grilles cartésiennes, une grille BCC offre un bon compromis entre robustesse, précision et coût calculatoire induit. Elle est particulièrement compacte et présente une faible anisotropie de voisinage, ce qui permet de réduire les erreurs de

métrification introduites par une optimisation discrète de la norme L_1. La résolution de la grille utilisée est fixée à $0,3$mm par voxel, de l'ordre de la résolution intra-coupe typique de nos données CT cardiaques. On peut enfin noter que puisque les nœuds du graphe correspondent aux positions des points de contrôle, les points centraux $p_t = \frac{c_t+c_{t-1}}{2}$ sont donc définis comme les centres des arcs (voir figure 7.7). Le voisinage définit par ailleurs la discrétisation des directions tangentielles possibles, au nombre de 14.

Les rayons sont discrétisés sur un intervalle uniforme $R = [\![r_{min}, r_{max}]\!]$, avec $r_{min} = 0,3mm$, $r_{max} = 3,9mm$ et un pas $s_r = 0.3mm$. Ces valeurs sont motivées par l'observation des rayons présents dans notre base de segmentations manuelles (voir figure 7.8) et découlent d'un compromis entre coût calculatoire et précision. En termes de voisinage, nous considérons initialement toutes les transitions possibles entre rayons successifs. La section 7.4 décrit comment restreindre dynamiquement ce voisinage.

Le graphe 4D final est illustré dans les figures 7.9 et 7.10. La structure de voisinage multi-échelle 4D est obtenue comme le produit du voisinage spatial et du voisinage de rayon. Le graphe est donné par $G = \{V, E\}$ avec les nœuds $V = \{(c_i, r_j)\}$ correspondant à des couples de points de contrôle et de valeurs de rayon. Les états $x_t = (p_t, r_t, d_t) \equiv (c_{t-1}, c_t, r_t)$ sont implicitement encodés par les arcs du graphe (voir figure 7.10), permettant la prise en compte d'une discrétisation grossière des orientations sans augmenter la dimensionalité du graphe (Gulsun and Tek, 2008b).

En comparaison à d'autres méthodes de chemins minimaux multi-échelles telles que (Wink et al., 2004; Li and Yezzi, 2007; Benmansour et al., 2009), la structure proposée diffère sensiblement. En particulier, notre structure de voisinage est plus riche et nos coûts cumulatifs sont exprimés sur les arcs plutôt que sur les noeuds. L'évaluation de ces coûts est détaillée dans la section 7.3.5. Nous expliquons notamment comment les différentes distributions de probabilités sont discrétisées, et comment les réponses de notre primitive `MFlux` sont évaluées le long des arcs du graphe (voir figure 7.9) de façon similaire à (Gulsun and Tek, 2008b) (voir figure 7.6).

1.6.4 Restriction de l'exploration de l'espace d'échelle

Le graphe 4D engendré par la discrétisation de notre espace de recherche est particulièrement grand et complexe. Dans la section 7.4, nous proposons deux stratégies heuristiques destinées à limiter les coûts de temps de calcul et de mémoire liés à la propagation itérative de chemins minimaux sur cette structure :

1. un élagage statique, a priori, des transitions de rayon possibles ;
2. un processus de sélection dynamique des réalisations / chemins propagé(e)s.

La restriction a priori des transitions de rayon possibles est réalisée en ne gardant que les N_r transitions les plus probables au sens de notre a priori $p(r_t|r_{t-1})$ (voir section 6.4.1 et figure 7.11). Cette stratégie correspond à un élagage du graphe, c'est-à-dire à la suppression de certains arcs. Dans nos expérimentations, nous n'avons observé quasiment aucune dégradation du résultat jusqu'à $N_r = 4$.

Notre seconde stratégie restreint dynamiquement les chemins propagés. Nous proposons de ne propager que les $H < |R|$ meilleures réalisations de rayons différents atteignant une certaine position spatiale c_i. Cette stratégie découle de l'observation que, pour une position spatiale donnée, l'ambiguïté relative au rayon est généralement faible. L'intervalle de rayons pertinents, autour du rayon réel, est classiquement restreint. Pour $H = |R|$, notre méthode correspond à une exploration complète du graphe 4D. Pour $H = 1$, elle

s'approche de méthodes mono-échelle telles que celle de (Gulsun and Tek, 2008b). Le paramètre H contrôle un compromis entre efficacité calculatoire et robustesse aux ambiguïtés relatives à l'échelle d'analyse. En pratique, des résultats satisfaisants sont obtenus en ne propageant qu'un nombre restreint de chemins par position spatiale, typiquement, $H = 4$ (voir section 7.7).

La combinaison de ces deux stratégies permet des gains appréciables en coût de calcul et de mémoire, rendant notre approche utilisable en pratique sur des ordinateurs commerciaux actuels.

1.6.5 Description algorithmique et implantation

La section 7.5 donne les détails d'algorithmie et d'implantation de notre approche. Comparée aux implémentations classiques de l'algorithme de Dijkstra, notre méthode introduit certaines subtilités, notamment pour la mise en place des stratégies d'élagage décrites précédemment.

La complexité (algorithmique) de notre algorithme, $O(NH\log(NH))$ pour un graphe de N nœuds, découle directement de celle de l'algorithme de Dijkstra. Cette complexité est donc du même ordre que celles d'autres approches multi-échelles de la littérature (Wink et al., 2004; Li and Yezzi, 2007; Benmansour et al., 2009) ($O(NR\log(NR))$, avec R le nombre de rayons considérés). La principale différence est que notre approche n'explore qu'un sous-ensemble restreint et dynamiquement adaptatif de la dimension liée au rayon ($H < R$), permettant des gains substantiels en temps de calcul et en consommation de mémoire.

Cette section décrit les structures de données employées pour obtenir de hauts niveaux de performance, que ce soit vis-à-vis du temps de calcul ou de la consommation mémoire. Les principales procédures algorithmiques composant notre approche sont également données. Les subtilités techniques présentées ont un impact de premier ordre sur les performances pratiques de l'algorithme, surtout quand celui-ci est appliquée à des données de grande taille telles que les acquisitions cardiaques en imagerie CT.

1.6.6 Schéma interactif et schéma automatique

La section 7.6 détaille deux scénarios d'utilisation de notre schéma d'extraction :

- une utilisation interactive, où l'optimisation est réalisée entre deux points fournis par l'utilisateur ;
- une utilisation automatique, où les arbres coronaires complets sont extraits entièrement sans interaction manuelle.

Scénario interactif Le scénario interactif correspond certainement à l'utilisation la plus classique de méthodes fondées sur l'optimisation de chemins minimaux. Pour notre application aux artères coronaires, nous faisons l'hypothèse que le point graine de début fourni par l'utilisateur correspond à une position proximale par rapport au point graine de fin. La propagation de chemins sur le graphe peut être stoppée dès lors que le point de fin est atteint. On peut noter que ce scénario n'est pas nécessairement limité à la spécification d'un unique point de fin. On peut ainsi spécifier plusieurs points distaux et extraire un arbre coronaire entier grâce à une unique procédure d'optimisation (voir figure 7.26).

Scénario automatique Le scénario d'utilisation automatique repose sur plusieurs étapes successives :

1. la segmentation automatique de l'aorte ;
2. la détection automatique des ostia ;
3. la propagation de chemins minimaux à partir des ostia, jusqu'à ce qu'un critère d'arrêt soit rempli ;
4. l'extraction, parmi tous les chemins propagés, des branches significatives des arbres coronaires.

La segmentation de l'aorte est effectuée grâce à un algorithme d'optimisation sur graphe développé à Siemens Corporate Research (Grady and Schwartz, 2006). Les ostia sont ensuite détectés à l'aide d'un processus de croissance illustré dans la figure 7.18. Ce type de schéma a précédemment démontré sa fiabilité dans d'autres travaux (Dikici et al., 2008; Tek et al., 2008; Gulsun and Tek, 2008b). Un exemple de résultat pour cette étape préliminaire est donné dans la figure 7.19.

Pour chaque ostium détecté, notre approche est utilisée pour propager des chemins explorant l'arbre vasculaire. Des masques de l'aorte et du cœur (Funka-Lea et al., 2006) sont exploités pour contraindre la zone de propagation (voir figure 7.20). La propagation est stoppée suivant un critère d'arrêt détaillé dans la section 7.6.2.2. Ce critère est défini comme un seuil sur le coût de mise à jour correspondant à la valeur au-delà de laquelle le chemin considéré est plus vraisemblablement en train de se propager dans des structures non vasculaires. Grâce à notre modèle de vraisemblance fondé sur une compétition entre vaisseaux et fond (voir section 6.3), ce seuil peut être aisément défini suivant un ratio de vraisemblances. Pour plus de robustesse aux arrêts prématurés (en cas de sténoses, de calcifications ou de baisse locale du contraste), le critère d'arrêt porte sur plusieurs arcs successifs.

A convergence, l'arbre vasculaire est extrait suivant un processus hiérarchique illustré dans la figure 7.21. Ce processus ordonne les chemins propagés en ordre décroissant de longueur. Ainsi, les branches les plus longues sont extraites en premier. L'information volumique correspondante est extraite et mise à jour progressivement sous forme d'un masque binaire. Grâce à ce masque, les chemins "parasites", c'est-à-dire les chemins correspondant à de faibles déviations de véritables branches vasculaires, peuvent être éliminés suivant le principe proposé par (Gulsun and Tek, 2008b). Ce processus hiérarchique élégant permet d'extraire l'arbre vasculaire et de détecter les bifurcations rétrospectivement. En comparaison, les approches par suivi direct doivent généralement gérer la délicate tâche de détecter les bifurcations en ligne, au moment même où elles sont rencontrées.

1.6.7 Evaluation

La section 7.7 propose une évaluation qualitative et quantitative de notre approche.

Evaluation qualitative La robustesse de notre méthode est notamment démontrée sur différents exemples cliniques, en présence de vaisseaux fins et à forte courbure, de calcifications, de sténoses, de stents et même d'occlusions totales (voir figures 7.25, 7.22, 7.23, 7.24). Les figures 7.26 et 7.27 donnent des exemples de résultats obtenus avec la variante automatique de notre approche. Celle-ci s'avère relativement robuste aux arrêts prématurés (faux négatifs) et problème de "fuite" (faux positifs). Parmi les erreurs les plus fréquemment rencontrées, on peut mentionner l'extraction erronée des extrémités allongées

de ventricules cardiaques et l'extraction de veines à proximité des artères coronaires d'intérêt (voir figure 7.28). Ces problèmes peuvent être atténués par l'utilisation d'un critère d'arrêt plus strict, au risque cependant d'arrêts prématurés plus fréquents.

En termes d'efficacité calculatoire, notre implantation s'avère particulièrement efficace. Dans un scénario interactif, une artère coronaire est en moyenne extraite en moins d'une minute pour $H = 4$ (moins de 15 secondes pour $H = 1$) sur un processeur `Intel Core Duo Intel` à 2.16GHz. Les temps d'exécution pour le scénario automatique restent du même ordre de grandeur. Le coût calculatoire dépend notamment directement de la fraction de l'espace de recherche qu'il est nécessaire d'explorer avant convergence (atteinte du point de fin ou critère d'arrêt).

Il est intéressant de noter l'effet bénéfique de l'utilisation d'une distribution de vraisemblance pour le fond, illustré dans la figure 7.29. Celle-ci permet en effet de réduire considérablement (jusqu'à 50%) l'étendue de l'exploration en pénalisant les chemins se propageant dans des zones non-vasculaires. Suivant le même raisonnement, notre modèle d'observation bi-partite a un effet bénéfique notable sur la robustesse aux phénomènes de décentrement et de raccourcis.

L'a priori sur les variations de rayon améliore la robustesse en présence d'anomalies et en cas de faible rapport signal sur bruit (figure 7.30). L'algorithme reste malgré tout capable de s'adapter à des variations relativement rapides de rayon (figure 7.25).

L'effet qualitatif du paramètre H (nombre de chemins propagés par position spatiale) est plus délicat à appréhender. Dans une majorité de cas, les résultats obtenus pour différentes valeurs de H sont très similaires, confirmant que l'ambiguïté liée au rayon est souvent très limitée. La figure 7.31 illustre cependant une situation où le chemin correct est extrait pour $H = 4$ quand $H = 1$ mène à un résultat erroné. L'évaluation quantitative proposée dans la section 7.7.2 confirme qu'augmenter H améliore la robustesse de l'algorithme. En revanche, nous n'avons pas observé d'amélioration significative au-delà de $H = 4$.

Evaluation quantitative Les deux scénarios d'utilsation (interactif et automatique) ont tout d'abord été évalués quantitativement sur notre base de segmentations manuelles (annexe B). Dix jeux de données furent sélectionnés aléatoirement pour l'apprentissage de notre modèle bayésien, laissant 51 jeux de données pour l'évaluation. Cette base de test totalise plus de 800 segments vasculaires, comprenant aussi bien des branches coronaires principales que de fines branches secondaires. Les tables 7.1 et 7.4 synthétisent les résultats de cette évaluation pour le scénario interactif et le scénario automatique, respectivement. Les critères employés sont détaillés au début de la section 7.7.2.1. La version interactive obtient d'excellents résultats en termes de robustesse (coefficient moyen de recouvrement avec la segmentation manuelle OV de 98,4% pour $H = 1$ et jusqu'à 99,3% pour $H = 4$). Les précisions de localisation de la ligne centrale et d'estimation du rayon (scores AI et AR, respectivement) sont d'ordre voxélique (la résolution typique de nos données est de l'ordre de $0,3$mm). Les résultats obtenus pour l'approche automatique sont, comme attendu, inférieurs, avec un taux de recouvrement moyen $OV = 85\%$. Ce score s'explique notamment par un plus haut taux de faux positifs (16.5%). Une statistique complémentaire intéressante est le taux OT de recouvrement pour les branches principales (diamètre supérieur à 1,5mm). Celui-ci atteint 92,9%, confirmant que l'algorithme est capable d'extraire automatiquement et avec fiabilité les segments coronaires principaux.

En complément à cette évaluation sur notre base de données interne, nous avons également évalué nos algorithmes sur la base d'évaluation du Rotterdam Erasmus Medical Cen-

ter (Rotterdam Coronary Artery Algorithm Evaluation Framework Schaap et al. (2009a)). Cette évaluation externe permet de rendre nos résultats directement comparables à ceux d'autres méthodes de la littérature, et d'évaluer les résultats comparativement à une variabilité inter-opérateurs. La base d'évaluation de Rotterdam comprend 24 jeux de données de test. Les tables 7.2 et 7.5 synthétisent les résultats obtenus pour les scénarios interactif et automatique, respectivement. Ceux-ci confirment l'excellente performance de la version interactive, avec un taux de recouvrement moyen OV de 97,6%. On peut notamment remarquer une amélioration sensible des performances comparativement à la méthode de Metz et al. (2008a) ($OV = 91,9\%$) également fondée sur l'extraction de chemins minimaux (voir la table 7.3 pour une comparaison des scores de différentes méthodes de la littérature). A ce jour, la seule méthode publiée et obtenant de meilleurs résultats est l'approche de Friman et al. (2008b) avec $OV = 98,5\%$. Celle-ci a néanmoins requis plus d'interaction que la nôtre en utilisant des points intermédiaires le long du vaisseau, là où nous nous sommes limités à l'utilisation des seuls points de début et de fin fournis. Notre version automatique obtient un taux de recouvrement de 91.3%, de l'ordre de la variabilité inter-opérateurs. Il est à noter que l'évaluation de Rotterdam se focalise sur les branches principales (4 par jeu de données) alors que notre évaluation interne portait sur les arbres coronaires entiers, ce qui explique les très bons résultats obtenus dans cette évaluation externe. Parmi les méthodes automatiques référencées (voir table 7.6), la seule obtenant à ce jour de meilleurs résultats vis-à-vis du recouvrement est celle de (Bauer and Bischof, 2008a) (92,7%), pour une précision légèrement inférieure. On peut également noter la supériorité de nos résultats (91.3% contre $84,7\%$) comparé à la méthode de (Tek et al., 2008), dérivée des travaux de (Gulsun and Tek, 2008b) et qui se fonde également sur une propagation automatique de chemins minimaux. Les principaux avantages de notre méthode sont l'utilisation d'une approche 4D optimisant conjointement la ligne centrale et le rayon et l'utilisation d'un a priori sur les variations de rayon dérivé de notre formulation bayésienne.

1.6.8 Discussion et perspectives

Le premier schéma d'extraction que nous proposons repose sur une optimisation discrète de chemins sur un graphe 4D (positions spatiales et rayon du vaisseau). Nos principales contributions résident dans la conception d'une nouvelle métrique cumulative, déduite empiriquement de notre modèle bayésien a posteriori. Celle-ci intègre notamment de l'information a priori sur les transitions de rayon le long des artères. Les différents termes la composant sont issus d'un apprentissage non-paramétrique sur une base de segmentations manuelles, évitant le caractère heuristique et paramétrique de nombreuses méthodes similaires. Nous avons également proposé plusieurs raffinements algorithmiques et d'implantation permettant d'atteindre en pratique de hautes performances en termes de temps de calcul et de consommation de mémoire. Enfin, notre évaluation qualitative et quantitative a permis de démontrer l'excellente robustesse de notre approche.

Parmi les perspectives d'extension et d'amélioration de cette approche, on peut citer l'implantation d'un schéma de propagation bi-directionnel pour accélérer les calculs (Deschamps and Cohen, 2001; Olabarriaga et al., 2003; Wink et al., 2000b) et l'utilisation potentielle d'heuristiques A^* permettant de réduire encore l'exploration de l'espace de recherche (Yuille and Coughlan, 1997). Le scénario d'utilisation automatique pourrait vraisemblablement être amélioré vers plus de robustesse aux faux négatifs en relaxant le critère d'arrêt via l'utilisation de techniques telles que le vote géodésique de (Rouchdy and

Cohen, 2008; Mille and Cohen, 2009). Enfin, les problèmes de faux positifs (phénomènes de "fuite") pourraient potentiellement être réduits grâce à des techniques d'élagage (Gulsun and Tek, 2008a) ou de post-traitement global (Bauer and Bischof, 2008a).

1.7 Chapitre 8 : Extraction par suivi stochastique (filtre particulaire)

Le chapitre 8 présente notre second schéma d'extraction, fondé sur un processus de suivi stochastique. Le principe de cette approche est d'estimer récursivement la distribution de probabilité a posteriori de notre modèle bayésien (chapitre 6). Elle exploite une technique dite de Monte-Carlo séquentielle, également connue sous le nom de filtre particulaire. L'estimation est réalisée par l'évolution d'une population d'échantillons aléatoires, les particules.

Dans un premier temps, nous motivons l'utilisation des filtres particulaires dans le cadre de la segmentation vasculaire et présentons les outils de base de ce type de méthodes. Nous détaillons ensuite notre implantation spécifique, qui inclut un certain nombre de raffinements tels qu'un schéma de propagation adaptatif couplé avec un processus de classification par algorithme de *mean-shift* pour la détection des bifurcations et l'extraction du résultat final.

Comparativement au schéma d'optimisation de chemins minimaux discrets (chapitre 7), le scénario d'utilisation de cette méthode se limite au suivi direct à partir d'un point de départ. Néanmoins, elle ne nécessite pas de discrétisation de l'espace de recherche et ne repose pas sur des schémas de sélection heuristiques. Notre évaluation qualitative et quantitative démontre le potentiel pratique de cette approche en termes de robustesse, précision et coût calculatoire.

1.7.1 Problématique et état de l'art

Nous sommes intéressés par l'estimation récursive de notre modèle bayésien présenté dans le chapitre 6. Son expression récursive générale est donnée par :

$$p(x_{0:t}|z_{1:t}) = \frac{p(x_t|x_{t-1})p(z_t|x_t)}{p(z_t|z_{1:t-1})}p(x_{0:t-1}|z_{1:t-1}) = \frac{p(x_t|x_{t-1})p(z_t|x_t)}{\int p(z_t|x_t)p(x_t|z_{1:t-1})dx_t}p(x_{0:t-1}|z_{1:t-1}) \tag{1.19}$$

où $x_t, t \in [\![0, L]\!]$, correspond à l'état de notre modèle et $z_t, t \in [\![0, L]\!]$, est l'ensemble des observations issues des données disponibles au temps t. Par hypothèse, les états suivent un processus markovien de premier ordre et les observations sont conditionnellement indépendantes étant donné l'état du modèle.

Dans le cadre du filtrage bayésien récursif, il est usuel d'exprimer cette récursion sous la forme d'une évolution en deux temps de la distribution marginale de filtrage $p(x_t|z_{1:t})$. La première étape est dite étape de prédiction :

$$p(x_t|z_{1:t-1}) = \int p(x_t|x_{t-1})p(x_{t-1}|z_{1:t-1})dx_t \tag{1.20}$$

Cette équation gouverne la transition de la distribution $p(x_{t-1}|z_{1:t-1})$ entre $t-1$ et t grâce notamment à l'a priori de transition $p(x_t|x_{t-1})$. La deuxième étape est celle de mise à jour, prenant en compte les nouvelles observations à disposition :

$$p(x_t|z_{1:t}) = \frac{p(z_t|x_t)p(x_t|z_{1:t-1})}{p(z_t|z_{t-1})} = \frac{p(z_t|x_t)p(x_t|z_{1:t-1})}{\int p(z_t|x_t)p(x_t|z_{1:t-1})dx_t} \tag{1.21}$$

Les équations précédentes sont simples, tout du moins en apparence. Les intégrales qu'elles impliquent ne peuvent en général pas être calculées analytiquement (Doucet et al., 2000). Parmi les exceptions notables, on peut citer le filtrage de Kalman pour les processus gaussiens linéaires (Kalman, 1960) et les méthodes par grille pour les espaces d'état discrets (Arulampalam et al., 2002). Malheureusement, de par son apprentissage non-paramétrique, notre modèle bayésien implique un modèle d'observation non-linéaire et une dynamique non-gaussienne. De plus, notre espace de recherche 4D est trop grand pour être discrétisé et traité efficacement par des méthodes par grille (Arulampalam et al., 2002) . Afin de gérer un tel processus non-linéaire, non-gaussien et à espace d'état continu, il est nécessaire d'utiliser des méthodes approchées. Les techniques candidates incluent les filtres de Kalman étendus (Arulampalam et al., 2002) reposant sur une linéarisation locale du processus. Malheureusement, ce type de techniques continue de représenter la distribution a posteriori par une gaussienne et n'est donc pas en mesure de gérer les distributions multi-modales. Dans notre contexte applicatif, cela implique que les filtres de Kalman étendus ne peuvent gérer les bifurcations vasculaires. Une autre solution, adoptée dans notre méthode, est l'utilisation de méthodes de Monte-Carlo séquentielles (Doucet et al., 2000; Arulampalam et al., 2002). Ces méthodes, connues également sous le nom de filtres particulaires, reposent sur l'évolution stochastique d'une population d'échantillons. Leur application aux problèmes de filtrage de distribution a posteriori peut être considérée comme un schéma de suivi stochastique multi-hypothèse. Leur application au suivi des artères coronaires en imagerie CT est illustrée dans la figure 8.1.

A notre connaissance, la première application des filtres particulaires au suivi de structures vasculaires 3D peut être trouvée dans (Florin et al., 2005, 2006). Ces travaux initiaux ont inspiré de récents développements (Schaap et al., 2007a,b; Lesage et al., 2008; Allen et al., 2008) qui diffèrent dans leurs cadres applicatifs, leurs modèles de transition et d'observation et dans leurs schémas numériques d'échantillonnage. Ces différents travaux ont démontré l'intérêt de ce cadre théorique pour la segmentation vasculaire, en particulier en termes de robustesse. Les filtres particulaires offrent un cadre particulièrement flexible, permettant d'accommoder des modélisations de relativement haut-niveau. Nous proposons de les appliquer à l'estimation de notre modèle bayésien complet.

1.7.2 Bases des méthodes de Monte-Carlo séquentielles

La section 8.2 décrit les bases sous-jacentes aux méthodes de Monte-Carlo séquentielles. Le lecteur peut également se référer à des ouvrages de référence tels que (Doucet et al., 2001) pour des discussions approfondies sur les fondements théoriques et les propriétés pratiques de ces méthodes.

Echantillonnage de Monte-Carlo parfait L'un des principes fondateurs des filtres particulaires est l'utilisation de schémas d'intégration de Monte-Carlo. Les intégrales qui ne peuvent être calculées analytiquement sont estimées via la simulation d'échantillons aléatoires $\{x^i_{0:t}\}_{i=1}^N$. Si l'on a la possibilité de simuler directement suivant $p(x_{0:t}|z_{1:t})$, l'échantillonnage de *Monte-Carlo parfait* permet d'obtenir une estimation empirique $\hat{p}(x_{0:t}|z_{1:t})$:

$$\hat{p}(x_{0:t}|z_{1:t}) = \frac{1}{N}\sum_{i=1}^{N}\delta(x_{0:t} - x^i_{0:t})$$

En d'autres termes, la distribution continue est approchée par une somme discrète d'échantillons discret (voir figure 8.2). En général, il n'est malheureusement pas possible de simuler

directement suivant $p(x_{0:t}|z_{1:t})$.

Echantillonnage par importance L'échantillonnage par *importance* propose alternativement de simuler suivant une distribution d'importance $q(x_{0:t}|z_{1:t})$ (aussi appelée distribution de proposition) approchant la distribution a posteriori cible (Geweke, 1989). Une étape de correction est nécessaire pour obtenir des échantillons suivant $p(x_{0:t}|z_{1:t})$ asymptotiquement. Cette correction prend la forme d'une pondération $w(x_{0:t})$ des échantillons, de façon à obtenir :

$$\hat{p}(x_{0:t}|z_{1:t}) = \sum_{i=1}^{N} \tilde{w}_t^i \delta(x_{0:t} - x_{0:t}^i)$$

Echantillonnage séquentiel par importance Les méthodes d'échantillonnage par importance peuvent être formulées sous une forme séquentielle, mieux adaptée aux problèmes récursifs. Cette reformulation impose certaines conditions sur la distribution d'importance utilisée (Doucet et al., 2001), qui s'exprime sous la forme :

$$q(x_{0:t}|z_{1:t}) = q(x_t|x_{0:t-1}, z_{1:t})q(x_{0:t-1}|z_{1:t-1})$$

En d'autres termes, la distribution d'importance est exprimée de façon récursive et comprend une composante de prédiction $q(x_t|x_{0:t-1}, z_{1:t})$. Le processus d'estimation peut désormais se focaliser uniquement sur la vraisemblance marginale $p(z_t|x_t^i)$ et l'a priori de transition $p(x_t^i|x_{t-1}^i)$, sans nécessité de modifier les échantillons passés. Le plus simple des filtres particulaires fondé sur ces principes est illustré dans la figure 8.4. Son efficacité pratique est en général limitée, notamment à cause de problèmes de dégénérescence.

Dégénérescence Il est prouvé que la variance de la pondération d'importance des échantillons augmente de manière exponentielle avec la longueur de la séquence (Gilks and Spiegelhalter, 1996). Cela signifie que pour une méthode d'échantillonnage séquentiel par importance, la population d'échantillons dégénère rapidement si bien qu'après quelques itérations, seulement un très faible nombre d'échantillons a un poids significatif (voir figure 8.6). En pratique, un algorithme de suivi fondé sur cette méthode risque de perdre très rapidement sa cible. Un estimateur de cette dégénérescence peut être trouvé dans la mesure du nombre d'échantillons effectifs (*effective sample size*, ESS).

Rééchantillonnage Les techniques de rééchantillonnage permettent de réduire les effets de dégénérescence en réallouant les échantillons plus efficacement. Leur principe général est d'obtenir des échantillons approximativement distribués suivant la distribution a posteriori cible à partir des échantillons pondérés obtenus par importance. Intuitivement, les échantillons de faible poids tendent à être détruits et ceux de fort poids à être dupliqués (voir figure 8.7). Il existe dans la littérature de nombreux schémas numérique de rééchantillonnage. L'une des techniques les plus populaires, pour sa simplicité et ses performances empiriques, est certainement le rééchantillonnage systématique (Arulampalam et al., 2002; Douc et al., 2005) décrit en figure 8.8.

En couplant échantillonnage séquentiel par importance et rééchantillonnage, on obtient l'algorithme SIR (*Sampling-Importance-Resampling*) aussi connu sous le nom de filtre *Bootstrap* (Gordon et al., 1993) ou *Condensation* (Isard and Blake, 1998b).

Appauvrissement de l'échantillonnage L'une des contreparties du rééchantillonnage est une perte de diversité dans la population d'échantillons, dès lors que les échantillons de fort poids sont dupliqués. Parmi les techniques proposées pour lutter contre ce phénomène, on peut citer l'utilisation de la régularisation par noyau pour perturber l'échantillonnage (Musso et al., 2001) ou l'utilisation de mouvements de type *Markov Chain Monte-Carlo* (MCMC) pour "rajeunir" la population (Gilks and Spiegelhalter, 1996; Gilks and Berzuini, 2001). Cette approche fut notamment appliquée à la segmentation vasculaire dans (Allen et al., 2008).

Choix de la distribution d'importance Le choix de la distribution d'importance est une étape importante dans la conception d'un filtre particulaire. Idéalement, cette distribution doit être proche de la distribution cible, c'est-à-dire la distribution a posteriori marginale. Le choix d'une distribution d'importance adéquate permet de réduire considérablement les problèmes de dégénérescence et d'appauvrissement.

Parmi les choix les plus populaires, on trouve l'utilisation de l'a priori de transition comme distribution d'importance : $q(x_t|x_{t-1}^i, z_t) = p(x_t|x_{t-1}^i)$ (Gordon et al., 1993; Isard and Blake, 1998b). Ce choix simplifie l'expression de la pondération d'importance, mais ne prend pas en compte les dernières observations z_t disponibles, ce qui peut devenir problématique en cas de désaccord important entre l'a priori et la vraisemblance.

Différents raffinements ont été proposés dans la littérature pour obtenir de meilleures distributions d'importance (Casella and Robert, 1996; Van der Merwe et al., 2001; Godsill and Clapp, 2001; Oudjane and Musso, 2000). Parmi ceux-ci, les techniques de filtres particulaires auxiliaires (*auxiliary particle filters* (Pitt and Shephard, 1999)) sont réutilisés dans nos travaux. Enfin, on peut citer l'un de nos travaux précédents (Lesage et al., 2008) (voir annexe E) proposant un schéma de rééchantillonnage particulier, fondé sur une approximation discrète de la distribution a posteriori.

1.7.3 Filtre particulaire adaptatif

Notre approche introduit certains raffinements, comparée à des implantations plus classiques de type SIR. Tout d'abord, notre algorithme repose sur un schéma adaptatif, l'échantillonnage *auxiliaire*, pour construire la distribution de proposition. Ensuite, il utilise un nombre d'échantillons adapté dynamiquement en fonction de la complexité de l'estimation.

Filtres particulaires auxiliaires Les filtres particulaires auxiliaires (Pitt and Shephard, 1999) reposent sur l'idée de construire empiriquement la distribution d'importance en prenant en compte les dernières observations disponibles. Ce type de techniques procède en deux étapes successives :

1. une population auxiliaire est *simulée* par prédiction (typiquement au travers de l'a priori de transition) ;
2. la population d'échantillons finale est tirée suivant la population auxiliaire.

La première étape de simulation permet d'améliorer l'allocation des échantillons finaux dans la seconde étape. Intuitivement, la population auxiliaire permet la construction empirique d'une distribution de proposition prenant en compte les toutes dernières observations disponibles. Comparativement à un algorithme SIR utilisant classiquement l'a priori de transition comme distribution de proposition, le schéma auxiliaire permet généralement de

réduire les problèmes de dégénérescence en diminuant la variance des poids d'importance. Une discussion détaillée du principe de filtrage auxiliaire est donnée en section 8.3.1.

Adaptation du nombre d'échantillons Une majorité des filtres particulaires de la littérature emploient un nombre N fixe d'échantillons. Le choix de N dépend notamment de la difficulté et de la dimensionalité du problème et résulte d'un compromis entre robustesse, précision et coût calculatoire. Les travaux de (Fox, 2003; Soto, 2005; Lanz, 2007; Koller and Fratkina, 1998; Straka and Simandl, 2004, 2006) proposent d'adapter dynamiquement le nombre d'échantillons. Le but de telles techniques est double : augmenter la qualité et la robustesse des estimations en allouant plus d'échantillons dans les situations compliquées, mais aussi potentiellement diminuer le coût calculatoire en allouant moins d'échantillons dans les situations "aisées". Parmi ces techniques, nous réutilisons la méthode de (Straka and Simandl, 2006), qui alloue un nombre d'échantillons défini en fonction d'un nombre d'échantillons effectifs (ESS) fixe. L'idée principale est d'atteindre une qualité d'estimation fixe (estimée par l'ESS) et de faire varier le nombre d'échantillons en fonction de la dégénérescence de la population. Le nombre d'échantillons effectifs est estimé en fonction de la variance des poids d'importance (voir section 8.3.2).

AAPF Notre approche repose sur une combinaison des principes de filtrage auxiliaire et d'adaptation dynamique du nombre d'échantillons (algorithme AAPF). L'idée principale de notre méthode est d'utiliser l'étape de simulation du filtrage auxiliaire pour estimer, à chaque itération, le nombre N_t d'échantillons à tirer. Notre algorithme, décrit en figure 8.10, combine les effets du filtrage auxiliaire et de l'adaptation du nombre d'échantillons. Dans les situations délicates, l'algorithme alloue un nombre plus élevé d'échantillons, qui sont de plus généralement mieux repartis grâce au filtrage auxiliaire. Ces effets sont démontrés dans les expériences de la section 8.5.

1.7.4 Algoritme AAPF pour la segmentation des coronaires en imagerie CT

Notre algorithme AAPF combine filtrage auxiliaire et adaptation dynamique du nombre d'échantillons. Il demeure générique et applicable à n'importe quel problème d'estimation bayésienne séquentielle. La section 8.4 décrit l'implantation spécifique de cet algorithme pour notre application de choix, la segmentation des artères coronaires en imagerie CT.

L'algorithme AAPF est utilisé pour estimer notre modèle bayésien complet (chapitre 6). La population d'échantillons créée par l'algorithme correspond à des réalisations d'états de notre modèle géométrique $\{x_t^i = (p_t^i, d_t^i, r_t^i)\}$ avec p_t^i la position centrale, d_t^i la direction tangentielle et r_t^i le rayon de l'état x_t^i.

L'algorithme est initialisé à partir d'un point proximal, typiquement un ostium détecté suivant la procédure décrite dans la section 7.6.2. Il procède séquentiellement au suivi de l'arbre coronaire comme illustré dans la figure 8.1, jusqu'à ce qu'un critère d'arrêt soit rempli.

Discrétisation des distributions de vraisemblance et d'a priori Les distributions composant notre modèle bayésien, vraisemblances comme distributions a priori, sont apprises de façon non-paramétrique sur une base de segmentations manuelles. Afin de pouvoir évaluer les termes de vraisemblance et tirer aléatoirement suivant les distributions a priori, il est nécessaire de discrétiser ces différentes distributions. Les paramètres de discrétisation

utilisés sont donnés dans la table 8.1. Ils correspondent à un compromis entre précision, efficacité calculatoire et consommation mémoire. On peut noter cependant que les différentes quantités (valeurs de rayon, de changements angulaires et de réponses de `MFlux`) sont discrétisées relativement finement. Par exemple, la discrétisation des valeurs de rayon est environ 10 fois plus fine que celle employée pour notre schéma d'optimisation par chemins minimaux (chapitre 7). Notre modèle de transition $p(x_t|x^i_{t-1}) = p(r_t|r^i_{t-1})p(d_t|d^i_{t-1}, r^i_{t-1})$ comprend un a priori sur les variations de rayon et sur les variations de direction tangentielle. La section 8.4.2 décrit la procédure utilisée pour prédire de nouveaux échantillons par échantillonnage multinomial des a priori discrétisés.

Détection des bifurcations par *Mean-Shift* Les bifurcations artérielles se manifestent typiquement par l'apparition de plusieurs modes dans la distribution a posteriori estimée. Les filtres particulaires sont en théorie capable de gérer des distributions multimodales, mais celles-ci posent différents problèmes pratiques. En particulier, à cause des phénomènes d'appauvrissement, les modes secondaires sont rapidement perdus (voir figure 8.11). On peut noter que notre algorithme AAPF améliore sensiblement la robustesse aux branchements, comparé à une approche SIR classique. Les branches secondaires sont tout de même perdues après un certain nombre d'itérations.

Suivant les observations de (Florin et al., 2005, 2006; Allen et al., 2008) notamment, nous nous reposons sur un schéma de détection explicite des branchements pour pallier cette difficulté. Là où Florin et al. (2005, 2006) et Allen et al. (2008) ont utilisé un algorithme de classification par K-moyennes, nous préférons une approche par *Mean-Shift*. La classification par *Mean-Shift* peut être interprétée en termes d'estimation de densité par méthode à noyaux (Fukunaga and Hostetler, 1975; Comaniciu and Meer, 2002), qui s'intègre parfaitement dans notre cadre théorique d'estimation statistique. Deuxièmement l'algorithme de *Mean-Shift* ne requiert pas d'initialisation et ne nécessite pas la spécification préalable du nombre de classes, deux avantages précieux comparativement aux K-moyennes.

Les fondements théoriques de l'algorithme *Mean-Shift* sont discutés en détail dans la section 8.4.3. *Mean-Shift* se fonde sur la détection, par optimisation itérative, des modes de densité d'une population discrète (voir figure 8.13). Cette densité est estimée par une méthode à noyaux. Pour la détection des bifurcations coronaires, nous appliquons, à chaque itération t, l'algorithme *Mean-Shift* sur les positions $\{p^i_t\}_{i=1}^{N_t}$ des échantillons $\{x^i_t\}_{i=1}^{N_t}$, pondérées par leurs poids d'importance. Intuitivement, l'algorithme cherche donc les modes d'une densité continue, estimée par une méthode à noyaux sur notre estimation discrète de la distribution a posteriori (voir figure 8.14). Dans notre implantation, nous adaptons dynamiquement le paramètre de bande passante utilisé par *Mean-Shift* en fonction de l'estimation courante du rayon du vaisseau principal. Nous proposons également un schéma d'échantillonnage parcimonieux pour réduire le coût calculatoire de cette approche.

Lorsque plusieurs modes sont détectés, le filtre AAPF est relancé pour chaque souspopulation d'échantillons correspondante. Ainsi, chaque branche vasculaire est suivie indépendamment.

Critère d'arrêt D'une manière similaire à notre approche par propagation de chemins minimaux (section 7.6.2.2), nous proposons un critère d'arrêt fondé sur les ratios de vraisemblance $\frac{p_v(y^i_t|r_t=r^i_t)}{p_{bg}(y^i_t)}$ constituant notre modèle d'observation. Si ce ratio est inférieur à 1, l'observation y^i_t provient plus vraisemblablement d'une structure non vasculaire (fond)

que d'une véritable artère. Pour plus de robustesse, nous appuyons ici notre analyse sur toute la population d'échantillons $\{x_t^i\}_{i=1}^{N_t}$ à l'itération t. Les détails de ce critère d'arrêt sont donnés en section 8.4.4.

Extraction du résultat par *Mean-Shift* Les générations successives de populations d'échantillons fournissent une estimation, au sens de Monte-Carlo, de la distribution a posteriori récursive de notre modèle bayésien. Afin d'extraire de cette distribution une représentation compacte de la ligne centrale et du rayon associé, une première possibilité est de sélectionner, à chaque itération, l'échantillon de plus haut poids comme approximation de la réalisation de plus grande probabilité (Florin et al., 2005, 2006). Une seconde possibilité est d'utiliser les estimations, au sens de Monte-Carlo, de la position axiale et du rayon (Doucet et al., 2001; Allen et al., 2008). Dans nos travaux, nous préférons une troisième solution, qui est de réutiliser les modes spatiaux estimés par *Mean-Shift* pour la détection des bifurcations. Ceux-ci correspondent en effet aux maxima locaux de densité de la distribution a posteriori (marginalisée sur la dimension spatiale). La régularisation induite par *Mean-Shift* assure des trajectoires plus lisses que si nous avions sélectionné uniquement le meilleur échantillon. D'un autre côté, notre approche repose sur une estimation plus locale et généralement plus précise que les estimateurs de Monte-Carlo correspondants. L'estimation du rayon associé à la position de ligne centrale repose également sur une méthode à noyaux (voir section 8.4.5).

1.7.5 Expériences et validation

La section 8.5 propose une série d'expériences permettant de valider notre approche qualitativement et quantitativement. Les figures 8.21 et 8.22 donnent des examples de résultats sur des données CT cliniques.

Nous comparons tout d'abord notre algorithme AAPF avec d'autres filtres particulaires classiques tels que l'algorithme SIR et le filtre particulaire auxiliaire (APF) en fonction du nombre moyen d'échantillons alloués. La figure 8.16 illustre notamment la pertinence de notre schéma avec un nombre d'échantillons adaptatif, dont les performances moyennes en termes de robustesse surpassent celles de l'algorithme APF sur lequel il se base. En tant que méthodes stochastiques, les filtres particulaires peuvent engendrer des résultats différents à chaque exécution (voir figure 8.17). L'expérience illustrée dans la figure 8.18 démontre que notre schéma adaptatif permet également une décroissance plus rapide de la variance des résultats en fonction du nombre d'échantillons, ce qui favorise donc des résultats stables pour un nombre relativement bas d'échantillons.

Nous avons évalué notre approche quantitativement sur notre base de segmentations manuelles (51 jeux de données CT cardiaques), suivant un protocole équivalent à celui utilisé pour notre premier schéma d'extraction (voir section 7.7.2). Les résultats de cette validation sont discutés dans la section 8.5.2. La table 8.2 synthétisent les résultats quantitatifs pour l'approche SIR, APF et notre algorithme AAPF. De façon cohérente, notre approche obtient de meilleurs résultats que les deux autres filtres particulaires, notamment en termes de robustesse (coefficient de recouvrement avec la segmentation manuelle $OV = 86,2\%$ contre 83% et $84,3\%$ pour SIR et APF, respectivement). Les améliorations de la précision de l'estimation de la ligne centrale et du rayon restent marginales. On peut également noter la baisse du taux de faux négatifs pour AAPF comparativement aux deux autres approches.

Comparée à notre premier schéma d'extraction fondé sur la propagation de chemins

minimaux (scénario automatique), notre approche par filtre particulaire parvient à de meilleurs résultats en termes de robustesse ($86,2\%$ contre 85% pour le coefficient de recouvrement OV). Ce résultat est imputable à un taux de faux positifs moins élevé ($12,3\%$ contre $16,5\%$). En pratique, notre filtre particulaire semble moins enclin à extraire de façon erronée les veines et les extrémités fines des ventricules. Cette observation peut en partie s'expliquer par l'utilisation, dans notre deuxième schéma, d'un a priori sur les variations de direction et par une plus faible sensibilité aux problèmes de "raccourcis". En revanche, notre algorithme AAPF souffre d'un taux de faux négatifs légèrement plus élevé que la méthode par chemins minimaux ($15,4\%$ contre $13,5\%$). On peut néanmoins mettre en exergue les résultats obtenus pour les branches coronaires principales (diamètre supérieur à $1,5mm$), qui sont à la fois élevés et cohérents ($OT = 92.5\%$ en moyenne, d'écart-type inter-éxécution de 1.5%).

D'un point de vue qualitatif, les résultats obtenus avec notre filtre particulaire sont particulièrement satisfaisants. L'estimation de la ligne centrale est régulière et suit précisément la segmentation manuelle même dans ses portions de forte courbure (voir figures 8.21 et 8.22). Contrairement à notre première approche par chemins minimaux, notre filtre particulaire ne nécessite pas une discrétisation grossière de l'espace de recherche. Quantitativement, on peut ainsi remarquer une amélioration de la précision de la ligne centrale (erreur AI passant de $0,25$ contre $0,31$mm, pour une résolution intra-coupe typique de nos données de l'ordre de $0,3$mm par voxel).

1.7.6 Discussion et perspectives

Le second schéma d'extraction que nous proposons permet l'estimation récursive de notre modèle bayésien complet, réalisée grâce à une méthode de Monte-Carlo séquentielle. Cette approche peut être considérée comme une méthode de suivi directe, stochastique et massivement multi-hypothèses. Notre méthode fut notamment motivée par les travaux de (Florin et al., 2005, 2006) qui démontrèrent l'applicabilité de telles techniques à la segmentation vasculaire. Ces travaux ont également inspiré un certain nombre de méthodes récentes (Schaap et al., 2007a,b; Allen et al., 2008). Vis-à-vis des travaux existants, notre contribution porte sur plusieurs points :

- notre méthode se fonde sur notre modèle bayésien, dont tous les termes sont appris de manière non-paramétrique sur une base de segmentations manuelles ;
- nous avons introduit un schéma d'échantillonnage adaptatif (AAPF) combinant les principes du filtrage auxiliaire et une adaptation dynamique du nombre d'échantillons alloués ;
- nous avons proposé plusieurs raffinements pour l'application à la segmentation vasculaire, tels que l'utilisation de l'algorithme *Mean-Shift* pour la détection des bifurcations et l'extraction du résultat final.

Une validation qualitative et quantitative a permis de démontrer que notre approche (AAPF) améliore la robustesse des filtres particulaires classiques tout en diminuant leur variabilité inter-exécution. Comparativement à notre schéma par chemins minimaux, notre filtre particulaire donne des résultats plus réguliers et plus précis. En revanche, il est moins souple en termes de scénarios d'utilisation et souffre d'un taux de faux positifs légèrement plus élevé.

Parmi les pistes théoriques futures, on peut mentionner l'utilisation d'une méthode de saut de modèle (Isard and Blake, 1998a) pour améliorer la capture des fines branches secondaires via la modélisation explicite des branchements asymétriques. On pourrait éga-

lement entreprendre l'extension de notre approche à une approche d'estimation par *lissage*, prenant en compte les observations futures pour plus de robustesse (Kitagawa, 1996; Hürzeler and Künsch, 1998; Doucet et al., 2001). Dans notre contexte applicatif, la source d'observations (l'image) est en effet statique et entièrement accessible. De telles approches sont néanmoins beaucoup plus coûteuses en temps de calcul que les approches par filtrage, et ce malgré de récentes avancées algorithmiques prometteuses (Klaas et al., 2006).

Enfin, nous mettons en exergue les performances calculatoires de notre implantation. Pour un nombre d'échantillons $N^* = 1000$, un arbre coronaire est extrait en moins de 4 minutes en moyenne (processeur `Intel Core Duo Intel` à $2,16$GHz). Ces performances sont donc encore en-deçà de celles de notre approche par chemins minimaux (moins de 1 minute). Néanmoins, notre filtre particulaire a l'avantage de pouvoir être aisément et massivement parallélisé. Nous pensons ainsi qu'une implantation sur carte graphique pourrait atteindre de très hauts niveaux de performance.

1.8 Chapitre 9 : Conclusions et perspectives

Au cours de ces travaux de thèse, nous avons discuté différents points fondamentaux pour la conception de méthodes de segmentation vasculaire, suivant trois axes d'étude : les modèles, les primitives et les schémas d'extraction. Cette démarche méthodologique nous a notamment apporté une meilleure appréhension de la vaste littérature du domaine. Nos développements ont porté sur une application médicale de premier ordre et particulièrement complexe : la segmentation des artères coronaires en imagerie tomodensitométrique 3D (CT).

Les contributions de nos travaux incluent :

- un état de l'art étendu de la segmentation vasculaire 3D, articulé suivant nos 3 axes d'étude ;
- l'introduction d'un modèle géométrique discret, simple et particulièrement compact ;
- le développement de primitives fondées sur le flux de gradient dotées d'un fort pouvoir discriminant et d'un faible coût calculatoire ;
- l'introduction d'un modèle bayésien récursif pour les artères coronaires en imagerie CT, dont tous les termes sont appris d'une manière non paramétrique sur une base de segmentations manuelles ;
- le développement d'un premier schéma d'extraction fondé sur l'optimisation de chemins minimaux 4D (ligne centrale et rayon) dont la métrique cumulative est dérivée de notre modèle bayésien ;
- le développement d'un second schéma d'extraction mettant en œuvre une méthode de suivi stochastique par filtre particulaire, introduisant notamment un nouveau schéma d'échantillonnage adaptatif.

Pour chaque composant et chaque méthode, des études de validation sur une grande quantité de données cliniques ont permis de démontrer qualitativement et quantitativement la pertinence de nos choix. Une attention particulière a également été portée sur la mise en œuvre pratique de nos méthodes, via l'introduction d'optimisations algorithmiques et d'implantation. Ainsi, nos algorithmes se fondent sur des approches relativement sophistiquées tout en maintenant de haut niveaux de performance calculatoire, les rendant utilisables sur des ordinateurs commerciaux.

Notre contribution applicative a porté principalement sur une tâche préliminaire mais cruciale, qui est la délinéation des structures vasculaires d'intérêt. Nos méthodes peuvent ainsi s'intégrer directement à certaines routines cliniques, par exemple pour l'aide à la vi-

sualisation et au diagnostic (voir figure 9.1). La fiabilité des résultats obtenus par nos schémas d'extraction peut également bénéficier à l'initialisation de traitements subséquents. Une extension naturelle serait une segmentation précise de la lumière vasculaire en vue d'applications de quantification des pathologies vasculaires. La figure 9.2 illustre par exemple des résultats préliminaires obtenus via l'application de la technique de (Gulsun and Tek, 2006, 2008b,a) pour une segmentation précise des sections vasculaires.

Finalement, notre cadre méthodologique est général et transférable à d'autres applications vasculaires. La figure 9.3 illustre par exemple un résultat préliminaire obtenu avec nos méthode d'optimisation de chemins minimaux sur une angiographie CT des vaisseaux périphériques. L'un des principaux obstacles reste la disponibilité de bases de segmentations manuelles pour les étapes d'apprentissage et de validation. La modélisation et l'exploitation d'informations spécifiques à l'application cible demeurent prépondérantes pour traiter des problèmes aussi complexes que ceux rencontrés dans le domaine de la segmentation vasculaire.

Chapter 2

Introduction and Context

Vascular diseases are among the most important public health problems in developed countries (World Health Organization, 2008), motivating the tremendous amount of research dedicated to vascular imaging. In this context, the segmentation of vascular structures is particularly valuable for diagnosis assistance, treatment and surgery planning. Isolating the vessels of interest from complex datasets is indeed a critical seminal step for their accurate visualization and for the quantification of pathologies. Unfortunately, most angiographic clinical routines still rely heavily on manual operations. Given the amount of data generated by modern 3D imaging modalities, such as computed tomography angiography (CTA) and magnetic resonance angiography (MRA), manual segmentation can quickly add up to hours of processing. Automatic and semi-automatic image processing tools aim at easing and speeding up reviewing tasks, reducing the amount of manual interaction, lowering inter-operator variability and providing quantitative information.

Vascular segmentation is a particularly specific and challenging problem, which motivated a large amount of past and on-going dedicated research. In the context of automatic image understanding, the contributions of this thesis lie at different theoretical and practical levels. We first focus on modeling issues, in other words, on the injection of prior knowledge with the aim of improving both the robustness and the accuracy of the segmentation. Second, we study the design of adequate vessel-dedicated features, *i.e.*, what type of image information can be exploited to detect vascular structures and how this information is extracted in practice. Third, we investigate overall extraction schemes, *i.e.*, optimization algorithms driving the segmentation process, combining models and image information. As an application of these methodological axes, this thesis focuses on a task of prime medical importance, the segmentation of coronary arteries from 3D cardiac computed tomography (CT) data.

This thesis work was co-directed by Isabelle Bloch and Elsa Angelini from Telecom ParisTech, Paris, France, in collaboration with Siemens Corporate Research, Princeton NJ, USA, under the supervision of Gareth Funka-Lea.

2.1 Medical Context

2.1.1 Coronary Arteries and Coronary Heart Disease

Coronary arteries are the blood vessels running along the heart surface and supplying the cardiac muscle (myocardium) with oxygen-rich blood. Coronary arteries branch off the ascending aorta at so-called *ostia* locations. One distinguishes two coronary trees (Fig.

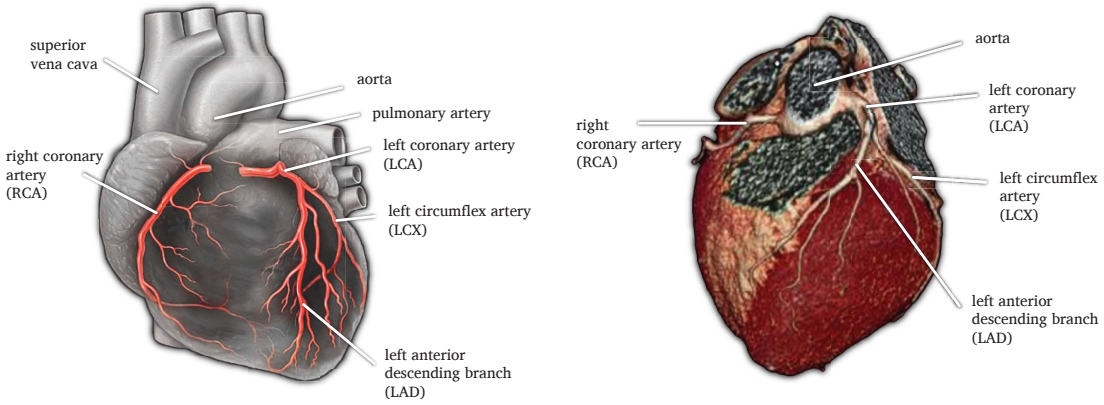


Figure 2.1: Anterior views of a human heart and principal vascular structures of interest. Left: stylized illustration (source: school of medicine, Yale university). Right: volume rendering of a cardiac CT angiogram (source: Siemens HealthCare).

2.1):

- the right coronary tree, whose main vascular branch is referred to as the right coronary artery (RCA);
- the left coronary tree, whose main vascular branch, the left coronary artery (LCA), separates between the left descending artery (LAD) and the left circumflex artery (LCX).

Coronary arteries play a critical role in the cardiac circulation system and are at the origin of severe medical conditions. Coronary heart disease (CHD) is currently the first cause of death in the US (Rosamond et al., 2008) and one of the leading cause of death worldwide (World Health Organization, 2008). CHD is caused by the accumulation of plaque in the arteries (atherosclerosis). Atherosclerosis yields a reduction of the coronary blood flow, resulting in conditions such as anginas and heart attacks. The decrease in quality of the myocardium perfusion can lead to heart failures and arrhythmias.

2.1.2 Computed Tomography for the Assessment of Coronary Heart Disease

Assessment of CHD is traditionally performed through invasive coronary angiographies (ICA) (Cademartiri et al., 2007), an X-ray procedure requiring the insertion of a catheter through the aorta for the injection of contrast agent (Fig. 2.2). Computed tomography angiography (CTA) proposes a non-invasive alternative to ICA (Mowatt et al., 2008). Following the injection of intra-venous contrast medium, cardiac CTA images consist of 3D acquisitions of the chest area. Modern multi-slice scanners reach sub-millimetric resolutions[1], enabling the accurate visualization of coronary arteries (Fig. 2.3). Whereas traditional ICA enables only the visualization of the vascular lumen, CTA acquisitions also reveal calcified and soft plaque (Fig. 2.4), which may be clinically relevant cues for risk assessment (Cademartiri et al., 2007; Schaap et al., 2009a).

Despite its advantages, the use of CT technology for coronary angiography raises sev-

1. Typical spatial resolutions for 64-slice CT scanners are currently of the order of $0.3 \times 0.3 \times 0.4$mm per voxel.

Figure 2.2: Traditional invasive coronary angiogram. Contrast agent (dye) is injected in the coronaries thanks a catheter, introduced through the aorta. Angiograms are obtained as 2D X-ray acquisitions. Source: Wikipedia.

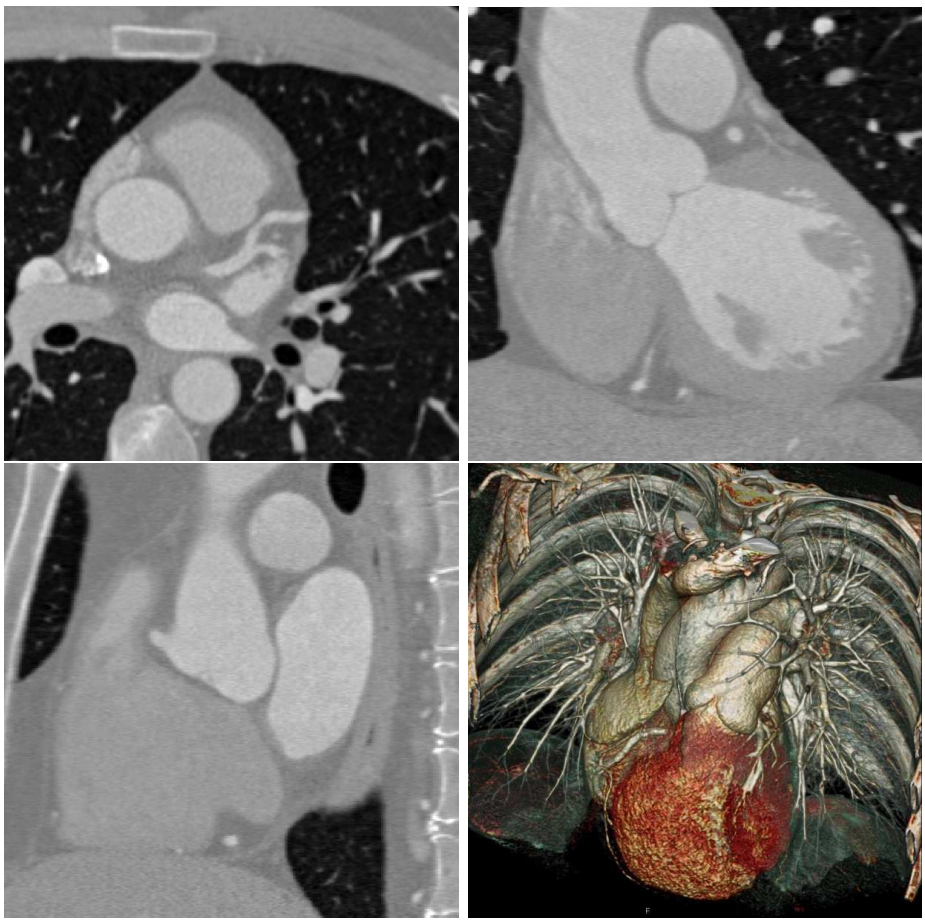


Figure 2.3: Cardiac computed tomography angiogram (CTA). Top, left: multi-planar reformation (MPR) axial view of a 2D slice through the 3D volume data. Top, right: coronal MPR view. Bottom, left: sagittal MPR view. Bottom, right: 3D volume rendering. Source: Siemens HealthCare, courtesy of Hong Kong Baptist Hospital Kowloon / Hong Kong, China.

Figure 2.4: Visualization of calcified and soft plaque in cardiac CTA data. Left: calcified plaque. Right: soft diffuse plaque. Curved planar reformation (CPR) views obtained from a delineation of the artery of interest. Source: Siemens HealthCare.

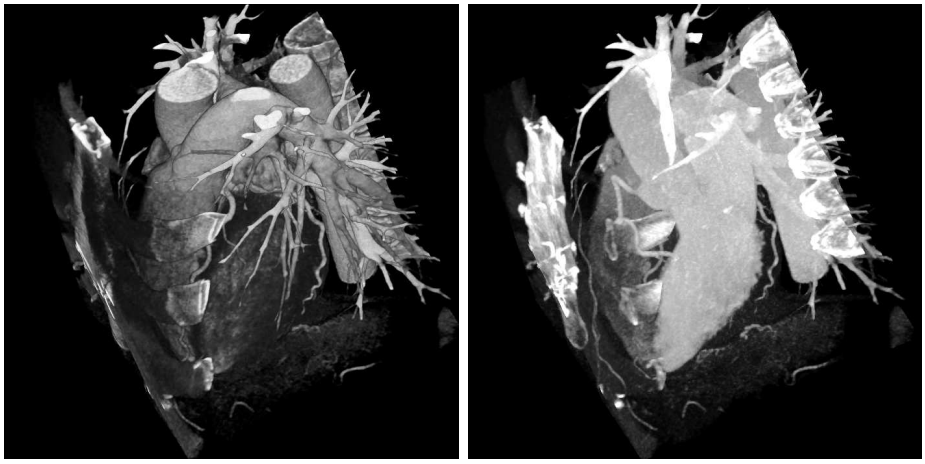


Figure 2.5: Left: volume rendering a cardiac CTA dataset. Right: maximum intensity projection (MIP) view of the same data.

eral concerns. First, CTA yields higher radiation doses than traditional ICA, and CTA acquisitions can suffer from motion, streaking (see Fig. 2.8, right) and partial volume artifacts. Modern multislice scanners help alleviate such issues by enabling faster acquisition times, without sacrificing the resolution and the field of view. Second, the amount of data in 3D CTA acquisitions is large and despite the high image resolution, coronary arteries are relatively small and thin structures (typical radiuses ranging from 10 to 1 voxel for segments of interest). This second issue makes manual review tasks potentially long and tedious.

Visual assessment of CHD from CTA data can typically be performed by scrolling through axial, coronal and sagittal multi-planar reformation (MPR) views (see Fig. 2.3). More advanced visualization techniques include curved planar reformation (CPR) and maximal intensity projection (MIP) views. Examples of CPR views are given in Fig. 2.4

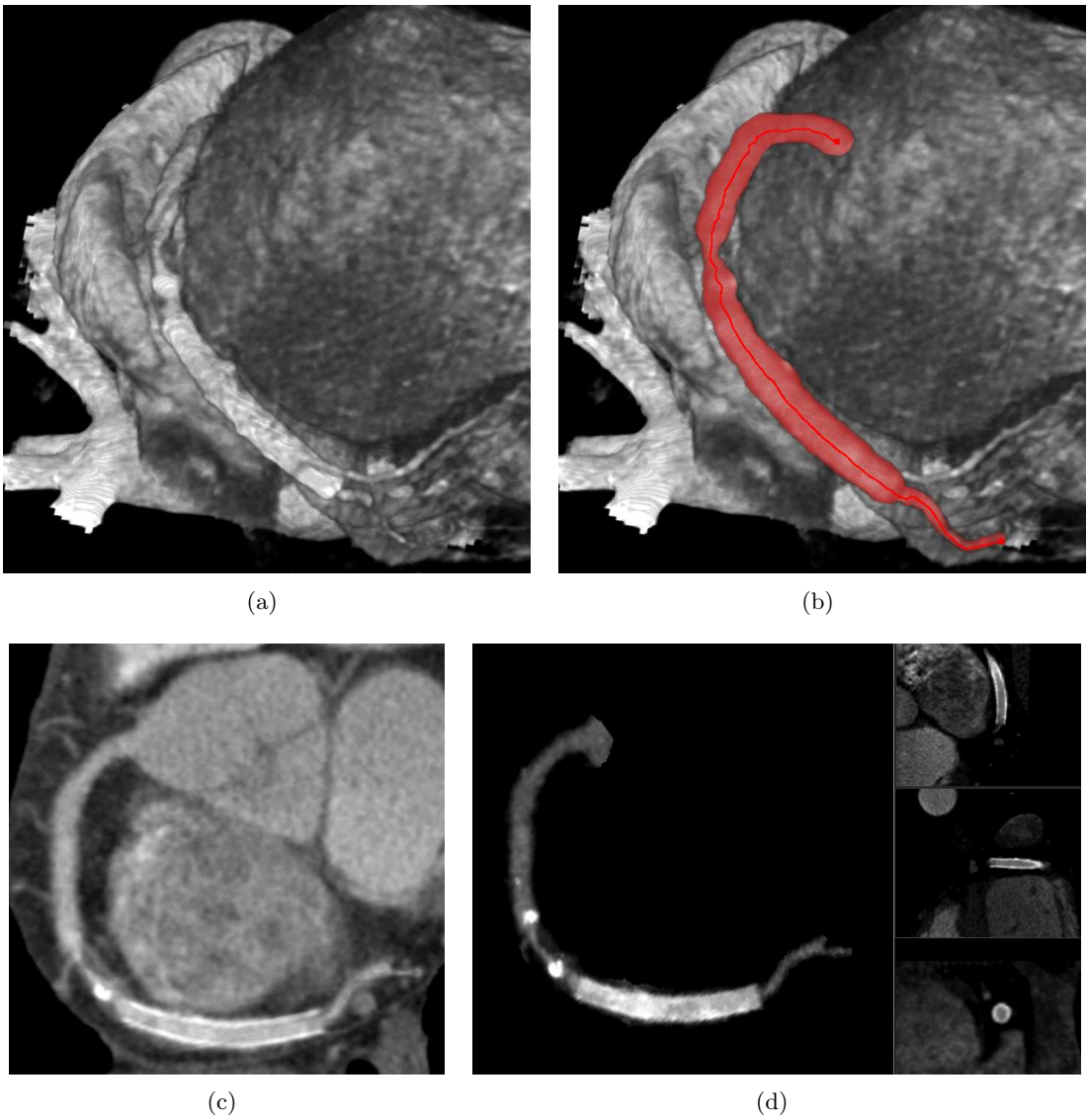


Figure 2.6: (a) Volume rendering of a right coronary artery with a stent, calcified plaque and stenosis. (b) Vessel extracted with the method from Chapter 7. (c) CPR view of the extracted branch. (d) Local tube MIP view of the extracted branch, with orthogonal MPR views of the stent.

and 2.6 (c). CPR techniques consist in generating 2D images where the entire length of the vessel of interest can be visualized, providing a global view of the target artery along with neighboring context. Different techniques can be employed to reconstruct CPR views (Kanitsar et al., 2002). Global MIP visualization of coronary arteries is made difficult by the presence of numerous hyper-intense structures other than coronary arteries: bones, pulmonary vessels and heart chambers in particular (Fig. 2.5). In practice, local MIP views can be obtained by constraining the projection in a small areas around a previously delineated vessel (Fig. 2.6 (d)).

CPR and local MIP visualization greatly simplify the visual assessment of potential

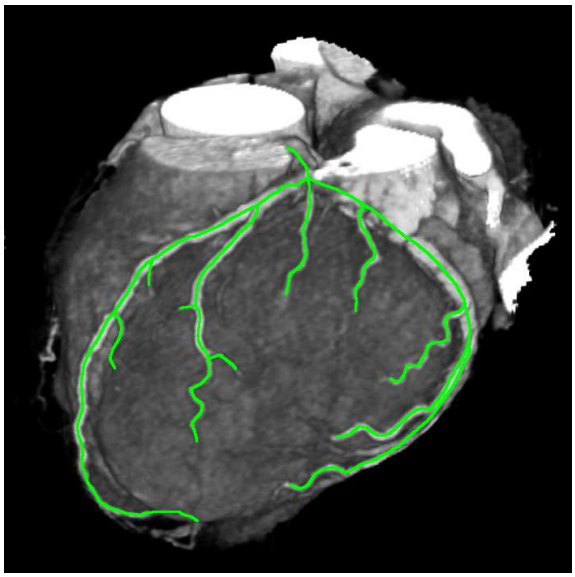


Figure 2.7: Sample of centerline delineation for a full left coronary tree.

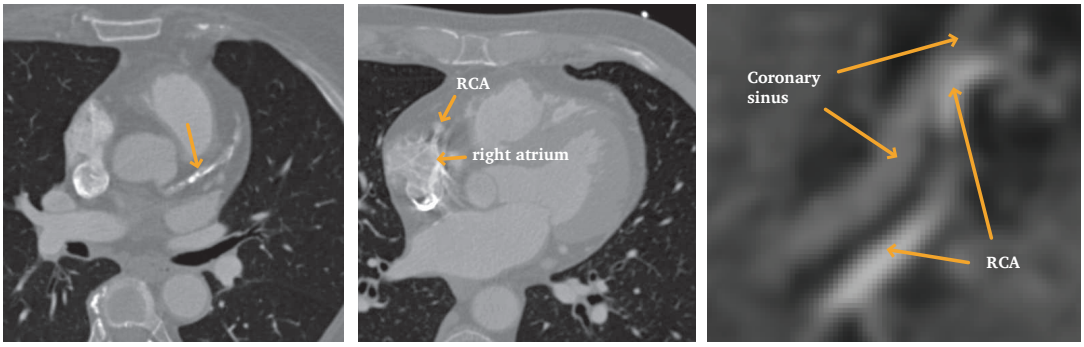


Figure 2.8: Image processing challenges raised by CT coronaries. Left: pathologies such as soft and calcified plaque altering the geometry and appearance of coronary arteries. Middle: spiking acquisition artifacts. Right: distal RCA branch running close to the coronary sinus. MPR views.

pathologies, but both require the preliminary delineation of the vessel(s) of interest. Automatic and semi-automatic techniques for the delineation of coronary arteries are thus particularly valuable in this clinical context. Last but not least, initial delineations can be exploited to greatly ease accurate lumen segmentation, enabling the automation of quantification tasks such as calcium scoring and stenosis grading.

2.2 Scope and Challenges of this Work

This thesis investigates the design issues encountered in the development of a vascular segmentation method through the case study of CTA coronary arteries. More particularly, we focus on the robust delineation of the vessels of interest. We consider delineation as the fundamental initial step of the segmentation process. From a reliable delineation of

the structures of interest, subsequent refinements, such as the accurate segmentation of the vessel surface, can be carried out in a more constrained and more robust fashion. Our objective is to extract a compact, high-level representation of the vascular network from the image data, in order to facilitate its subsequent interpretation and manipulation. In this context, the delineation of coronary arteries from cardiac CTA is a particularly challenging application. From an image processing point of view, difficulties come from various sources:

- the size of the data: cardiac CTA datasets typically reach sizes of the order of 512^3 voxels;
- the size of the structures of interest: conversely, coronary arteries are relatively thin, with radiuses ranging typically from 10 to 1 voxel in such data;
- acquisition-dependent issues: noise, partial volume effect, contrast inhomogeneities and artifacts (see Fig. 2.8, middle);
- the variability and complexity of coronary arteries: coronaries vary widely in sizes and curvature; they form complex vascular networks with numerous branches and high inter-patient variability (Fig. 2.7);
- pathologies (soft and calcified plaque), along with stents and bypasses altering the geometry and appearance of the coronaries;
- coronaries are surrounded with various hyper-intense structures of similar appearance: coronaries often run close to blood-filled heart chambers and vessels from the venous network (Fig. 2.8, right).

One important aspect of this work was the development of tools usable in a large array of realistic clinical scenarios. Motivated by the industrial collaboration with Siemens Corporate Research, special attention was thus given to the robustness and computational efficiency of the approaches we propose. We detail algorithmic and implementation contributions and evaluate our work both qualitatively and quantitatively on a large and diverse database of clinical datasets.

2.3 Outline

This document is organized as follows. In Chapter 3, we give an extensive review of state-of-the-art literature on 3D vascular segmentation techniques. Our bibliographic study is organized according to three methodological axes: geometric and appearance models, features and extraction schemes. We discuss the general trends in the literature, their theoretical and practical properties. Following the conclusions of our bibliographic analysis, Chapters 4 to 8 present original contributions of this thesis, dedicated to the segmentation of the coronary arteries from cardiac CTA data.

In Chapter 4, we introduce the discrete, medial-based representation we use to model the geometry of the vessels of interest. We additionally discuss the photometric properties of coronaries in CTA data, such as expected intensities and radial appearance profiles.

In Chapter 5, we present our multiscale, oriented vascular-dedicated feature, referred to as `MFlux` . The `MFlux` feature is used to evaluate the medial position, the orientation and the radius of coronary arteries. It is derived as a modification of a seminal, gradient-flux-based feature, referred to as `Flux` . The theoretical and practical properties of both `Flux` and `MFlux` features are discussed through a series of experiments and compared with a selection of features from the literature.

In Chapter 6, we introduce our Bayesian statistical framework to encode a probabilistic vessel model for CTA coronaries. This model is built upon our geometric model and our

feature. It combines model-based and data-based information into an *a posteriori* formulation. We detail notably how we learn prior knowledge on typical radius and direction transitions along coronary arteries and how we derive our likelihood observation model from distributions of `MFlux` responses.

In this thesis, we propose two different extraction methods, both relying on the aforementioned models and feature. Our first strategy, based on a minimal path optimization scheme, is presented in Chapter 7. This approach propagates realizations of our geometric model as discrete paths on a 4D (3D space+radius) graph. Paths are optimized with respect to a cumulative cost metric derived from our seminal Bayesian formulation. We discuss key theoretical and practical design issues and put our approach in perspective with related developments in the literature. We detail the implementation of both interactive and fully automatic workflows based on this minimal path formulation and present an extensive qualitative and quantitative evaluation study.

Our second extraction strategy, discussed in Chapter 8, is a stochastic, multi-hypothesis tracking approach based on particle filtering. We use particle filtering technique to estimate the successive states of our Bayesian model given the data. We introduce a number of algorithmic refinements specific to our applicative context and present a quantitative evaluation of our implementation for the automatic extraction of complete coronary trees.

Chapter 3

A Review of 3D Vessel Lumen Segmentation Techniques: Models, Features and Extraction Schemes

This chapter is an adapted and updated version of (Lesage et al., 2009a), published in Medical Image Analysis.

Angiography covers a wide range of anatomical applications (cerebral, retinal, hepatic, peripheral, pulmonary, cardiac...) and modalities (X-ray,CTA, MRA, ultrasound, 2D or 3D...). While not focusing on a particular application, we mainly report recent works on CTA and MRA in 3D. These modern modalities have reached high quality and resolution levels [1]. They allow the capture of highly detailed vascular networks embedded in increasingly large volumes of data. These two challenging aspects (complexity of the target network *v.s.* large search spaces) motivate the use of advanced, dedicated segmentation schemes.

Our bibliography study deals almost exclusively with the most common vascular segmentation task, *lumen* segmentation. Although most of the general points discussed in this chapter also apply to other vascular applications, tasks such as outer wall or thrombus segmentation raise additional, specific challenges.

In this review, we intend to remain as general as possible. We often do not detail application- and modality-dependent aspects. Instead, we focus on high-level discussions and emphasize general trends and major innovations in state-of-the-art literature. Our goal is to provide the reader with general insight on this particularly prolific field and highlight the key design issues that researchers may encounter in the development of new methods.

Given the wide range of applications considered, direct performance comparisons are also left out. More generally, one can regret a lack of standardly accepted databases and validation criteria for most vascular segmentation applications. Initiatives such as the BrainWeb project (Broche et al., 2006) for brain MRI and the Rotterdam Coronary Artery Algorithm Evaluation Framework (Metz et al., 2008b; Schaap et al., 2009a) for coronary CTA segmentation are particularly precious as they allow direct quantitative comparisons

1. Cardiac CT angiography, for instance, typically reaches sub-millimetric intra- and inter-slice resolutions with moderate anisotropy, on state-of-the-art multi-slice CT scanners.

between different approaches.

To the best of our knowledge, the currently most general and extensive vascular segmentation review can be found in (Kirbas and Quek, 2004), with a highly detailed categorization of existing works with respect to their mathematical frameworks. Other existing reviews focus more closely on a limited number of works and applications: CT peripheral angiography in (Felkel et al., 2001), MR angiography in (Suri et al., 2002a,b), with a first part (Suri et al., 2002a) dedicated to pre-filtering and a second part (Suri et al., 2002b) focusing on skeleton versus non-skeleton extraction schemes. One should additionally mention the review of skeletonization algorithms from (Buhler et al., 2002), which also covers visualization techniques, from surface and volume reconstruction, to the generation of meshes and curved planar reformation views.

Instead of using a linear categorization of the different frameworks, we discuss existing works according to three high-level axes:

1. appearance and geometric *models* (Sec. 3.1);
2. image *features* (Sec. 3.2);
3. extraction schemes (Sec. 3.3).

Models correspond to the prior assumptions made on the target vessels, *e.g.* elongation and hyper-intensity. *Features* are the vessel-dedicated image measures used to estimate the models on the image, *e.g.* local intensity curvatures. Finally, the *extraction scheme* represents the algorithmic core of a vascular segmentation method, *i.e.*, how models, guided by the features, are optimized to perform the segmentation in practice.

A given technique may rely on intricate combinations of models, features and extraction schemes. Table 3.1 illustrates the decomposition of a selection of works representative of the main trends in the field. Our primary goal is not so much the exhaustive categorization of existing works but the discussion of why proposed combinations are relevant. Each component effects the robustness, accuracy and efficiency of the final algorithm. Throughout this bibliographic study, we review the theoretical and practical properties of a large panel of models, features and extraction schemes.

Method	Models	Features	Extraction
Supervised wave propagation (Lorenz et al., 2003)	= hyper-intensity = locally tubular geometry = centerline straightness = branching frequency	= adaptive threshold – covariance analysis of extracted segments – front connectivity for branch detection	– wave propagation (Sec. 3.3.2) via fast-marching (one seed point) – hierarchical design with 3 levels (voxel, segment, tree) – various adaptation heuristics
Curves (Lorigo et al., 2001) (Fig. 3.17)	= 1D centerline curve – centerline smoothness from curvature regularization – varying, limited vessel radius	– image gradients	– geodesic, mean-curvature flow level-set evolution (Sec. 3.3.3) = co-dimension 2 (1D curve in 3D) = initialization from seed regions = implementation details (c-level set, reinitialization)
Model-Fitting Tracking (Wörz and Rohr, 2007, 2008)	– local 3D *hybrid* models (Sec. 3.1.3, Fig. 3.3): – Gaussian line, Phi line, bar-like convolved – centerline smoothness from Kalman filtering	– local fit of the hybrid model (Sec. 3.2.2)	– direct centerline tracking from a seed point (Sec. 3.3.4): – position prediction by Kalman filtering – correction by fit of the local model
Cores tracking (Fridman et al., 2003; Fridman, 2004) (extension of (Aylward et al., 1996; Aylward and Bullitt, 2002))	– intensity height ridge (Sec. 3.1.3): – hyper-intensity – local tubularity	– *medialness* cross-section-based feature (Sec. 3.2.3, Fig. 3.9) – corner detector + specific medialness pattern for bifurcation detection	– direct centerline tracking from a seed point (Sec. 3.3.4): – prediction by adaptive stepping – correction by local optimization of the medialness response – bifurcation detection
Hessian-based Minimal paths (Olabarriaga et al., 2003; Wink et al., 2000b, 2001, 2004)	= 1D centerline curve = ellipsoid model from Hessian-based features (Fig. 3.4, Sec. 3.1.3)	– multiscale Hessian-based vessel enhancement filters from (Frangi et al., 1998; Lorenz et al., 1997; Sato et al., 1998b) (Fig. 3.7, Sec. 3.2.2)	– L_1, Dijstra-like graph-based path minimization between two points (Sec. 3.3.4) – refinements for robustness and speed: – *pioneer points* (Wink et al., 2001) – bi-directional search (Olabarriaga et al., 2003; Wink et al., 2000b) – A* heuristic (Wink et al., 2000b)
4D minimal paths (Li and Yezzi, 2006) (Fig. 3.25)	= 4D paths: (centerline + radius) = circular cross-sections (inscribed spheres) = smoothness of radius and centerline variations from L_2 path minimization – 2 appearance models proposed: – similarity-based – contrast-based	– 2 intensity-based features proposed: – similarity to the seed sphere – contrast-based term over successive scales	– 4D fast-marching between start and end spheres (Sec. 3.3.4)
Particle filtering (Florin et al., 2005, 2006) (Fig. 3.26, Sec. 3.3.5)	= 1D centerline curve – cross-section geometry model: – elliptic in (Florin et al., 2005) – circular shortest paths in (Florin et al., 2006) = Gaussian mixture appearance = local hyper-intensity / contrast = *non-linear* variations of geometric and appearance parameters	– cross-sectional features: – contrast: *Ribbon* measure (Sec. 3.2.3) – Gaussian mixture estimation and K-L distance from the model	– multi-hypothesis, stochastic Bayesian tracking via particle filtering (Sec. 3.3.5) = direct tracking from seed point – clustering of particles for bifurcation detection
Markov marked point processes (Lacoste et al., 2006) (Sec. 3.3.5, Fig. 3.27)	– "Quality Candy" segment model: – locally linear segments of varying width and length – intensity homogeneity within segments – local contrast – connectivity of the segments – small curvature between consecutive segments – penalization of nearby parallel segments – Polyline refined tree model: – connectivity / influence between segments – local contour contrast	= contrast operator on linear segments = refined edge estimation (second stage)	– Markov marked point process: – simulated annealing of a Reversible Jump Monte Carlo Markov Chain process – reversible moves to perturb the segment population – combination of stochastic processes: – initial segment process – refined polyline hierarchical process – fully automatic, random initialization

Table 3.1: Sets of models, features and extraction schemes decomposed for a selection of illustrative works.

3.1 Vascular Models

Models, in a broad sense, embed prior information about the target structures. The principal distinction we make here is between appearance (photometric) and geometric information.

3.1.1 Appearance Models

Appearance models express prior knowledge on the luminance properties of the target vascular structures. The appearance of blood vessels in medical acquisitions is of course highly dependent on the imaging modality. For the sake of simplicity, we focus here on contrast-enhanced modalities (CTA and MRA) where vessels are brighter than the surrounding structures [2].

Vessel-Only Models Knowledge on the theoretical vessel intensity distribution can be derived from the specificities of the imaging protocols. For instance, CT angiographies are generally calibrated so that contrast-enhanced tissues appear in pre-defined intensity ranges. This property was directly applied for vessel segmentation in numerous papers (Boskamp et al., 2004; Sebbe et al., 2003; Wesarg and Firle, 2004; Zahlten, 1995). Detailed MR-specific appearance modeling aspects are discussed for instance in (Hoogeveen et al., 1998) for time-of-flight (TOF) and phase contrast (PC) acquisitions.

In practice, perfect steadiness of the contrast agent concentration is impossible to achieve and contrast agent inhomogeneity causes appearance variability, even for precisely calibrated modalities such as CTA. Such variability can be handled through the use of statistical distributions, derived from theoretical and/or heuristic knowledge. Various models have been proposed in the literature, such as Gaussian (Florin et al., 2005), Cauchy (Agam and Wu, 2005), Rician (Chung and Noble, 1999) or double sigmoid (Quek and Kirbas, 2001) distributions.

Vessels + Background Models When information on surrounding tissues is available, one can use a coupled appearance model, mixing hypotheses on the vascular intensity distribution and surrounding tissues. The simplest example is the assumption that vessels are brighter than their direct neighborhood in contrast-enhanced medical images, an hypothesis exploited in (Florin et al., 2005; Schaap et al., 2007b) among many others. A piece-wise constant model is used in (Tyrrell et al., 2007), with supposedly constant object and background intensities [3]. In the work of (Hoogeveen et al., 1998) dealing with MR angiography, the background steady state intensity is also modeled as constant. More advanced models are based on statistical mixtures (Chung et al., 2004; Florin et al., 2005; Gan et al., 2004; Hassouna et al., 2003; Manniesing and Niessen, 2004; Wilson and Noble, 1999; Yang et al., 2004). In numerous applications, vessel surroundings contain various structures of highly variable appearance, making the use of explicit background models more difficult. Instead of directly modeling background intensities, the work of (Schaap et al., 2007a) on carotid CTA segmentation exploits the assumption that vessels are actually more *homogeneous* than their surrounding structures.

2. In certain types of angiographic acquisitions, such as *black blood* MRA, vessels can also appear darker than their surroundings.

3. In (Tyrrell et al., 2007), authors actually work on 3D multiphoton images, where vessels are darker than the background.

Noise and Point-Spread Function Besides the appearance of the target vessels and background, noise is another acquisition-dependent aspect that can be modeled, estimated and incorporated as part of an appearance model. Uniform noise estimation was used in (Hassouna et al., 2003; Wilson and Noble, 1999), Rayleigh noise in (Hassouna et al., 2003) and Laplacian noise in (Tyrrell et al., 2007).

The *point spread function* (PSF) is another important characteristic of the imaging device and scan and reconstruction parameters. It accounts for spatial blurring phenomena observed in the image data. Explicit modeling of the PSF (generally as a spatial Gaussian function) was proposed in (Boskamp et al., 2004; La Cruz et al., 2004; Reinhardt et al., 1997; Saba et al., 2003) for instance. More generally, numerous *hybrid* models (see Sec. 3.1.3) and features (see Sec. 3.2.2 and 3.2.3) include such notions of spatial blurring inherent to the acquisition process. From a theoretical point of view, and although it is generally implicit, one can thus consider that such techniques integrate a model of the underlying PSF.

3.1.2 Geometric Models

A key characteristic of blood vessels is their specific shape. Their elongation, in particular, is valuable prior knowledge which can be encoded through geometric models. Depending on the application, additional knowledge (specific vessel radius ranges and bifurcations for instance) might be incorporated. We will discuss a few examples in Sec. 3.1.5. Hereafter, we focus solely on purely geometric models suited for most angiographic applications.

Surface Models Numerous works rely on implicit assumptions on the vessel surface, such as local tubularity. Such assumptions are generally used to model the cross-sectional geometric and appearance properties (see thereafter and Sec. 3.1.3). Among techniques exploiting directly a surface model, one can cite (Frangi et al., 1999a, 2000) with B-spline tensor surfaces and the works of (de Bruijne et al., 2003b,a) with *active shape models*, learned as shape distributions of abdominal aortic aneurysms (Fig. 3.1). More generally, one may consider that active contour techniques (see Sec. 3.3.3) incorporate model-based regularity assumptions on the vessel surface through the evolution forces they employ.

Vessels as 1D Curves: the Centerline A vessel can be considered as a primarily 1D structure, defined by its *centerline*, *i.e.*, the curve centered inside the vessel lumen. In terms of medial-based representations (Blum, 1967), a vessel's *skeleton* model simply reduces to a 1D curve. Owing to its simplicity and descriptive power, the vascular centerline has become a critically important tool, used in various clinical routines such as visual inspection tasks and stenosis or aneurysm quantification.

A classical assumption is that the centerline exhibits limited curvature. Various parametric models have been proposed to approximate smoothly varying curves: piece-wise linear curves in (Lacoste et al., 2006) and *B-splines* in (Bouix et al., 2005; Flasque et al., 2001; Frangi et al., 2000; Larralde et al., 2003). In (Wong and Chung, 2006), *cardinal splines* were preferred for their intuitive geometrical behavior. Assumptions on the centerline regularity can also be embedded in the extraction process using prediction techniques such as Kalman filtering (Kalman, 1960; Welch and Bishop, 2001) or particle filters (Arulampalam et al., 2002; Doucet et al., 2001). See Sec. 3.3.4 and 3.3.5, respectively, for further discussions.

Centerline+Cross-Section models, Generalized Cylinders A classical approach to enrich centerline models with contour information can be found in the family of *generalized cylinder* methods (Binford, 1971), by the use of *swept* surfaces. In such methods, the vessel surface is defined by sweeping 2D cross-sectional contours along the centerline. Proposed cross-sectional models include circles, ellipses, star patterns (O'Donnell et al., 1998; Verdonck et al., 1995; Williams et al., 1997) or parametric curves (Li and Ourselin, 2003). Such generalized cylinder models differ in the types of centerlines and cross-sections they can handle and in the way they maintain the 3D coherence of the structure. For instance, the *straight homogeneous generalized cylinder* model (Ponce et al., 1989) is limited to linear centerlines and cross-sections varying as scaled versions of a reference shape. Numerous extensions have been proposed (Azencot and Orkisz, 2003; Flórez Valencia et al., 2006; Huang and Stockman, 1993; Terzopoulos et al., 1988; Zerroug and Nevatia, 1996) for increased flexibility. Of particular interest, the *extruded generalized cylinder* from (O'Donnell et al., 1994) handles curved centerlines and varying cross-sections while preventing torsion artifacts common to previous models.

The technique from (Frangi et al., 1999a, 2000) relies, for its initialization, on a B-spline centerline curve. A coarse approximation of the vessel surface is then obtained from a swept generalized cylinder, which is subsequently refined by the optimization of a tensor-product B-spline surface (see Fig. 3.2). Similarly, a 3D active surface under axial constraints is proposed in (Montagnat, 1999) and a level-set evolution from an initial centerline is used in (van Bemmel et al., 2003). Yim et al. (2001) rely on a tubular coordinate system to evolve a parametric mesh with respect to the centerline curve, with a particular focus on parameterization schemes avoiding self-intersection issues. The recent works of (Mille et al., 2008; Mille and Cohen, 2009) also fall in the category of generalized cylinder methods, with the development of an explicit deformable model based on a centerline curve with varying associated radius. This model includes constraints on the smoothness of curve and radius variations, expressed as evolution forces.

A Contrario **Models** An original approach consists in describing a vessel *a contrario*, by defining what it should not be. Examples include the *ball measure* from (Nain et al., 2004) and the non-local measures used in (Rochery et al., 2005a, 2006). These *soft shape priors* are used in active contour frameworks to penalize local widening, maintaining the shape elongation *a contrario* (see Sec. 3.3.3 and Fig. 3.18 and 3.19). As it is generally inextricable to model every non-vessel structure present in the image, we expect such approaches to be limited to relatively low-level, soft geometric constraints.

3.1.3 Hybrid Models: Appearance + Geometry

Hybrid models combine appearance and geometric information by incorporating assumptions on the *spatial* appearance of a vessel.

Radial Intensity Profiles It is generally too complex to define the intensity distribution as a function of space for highly variable shapes such as generalized cylinders (see Sec. 3.1.2). The geometric part of hybrid models is thus usually limited to a *local* approximation of the vessel. The spatial intensity distribution is then modeled as radial variations in the corresponding cross-sections, introducing the notion of intensity *profiles*.

Mathematically appealing profiles include *parabolic* profiles (Steger, 1996) and *3D Gaussian line* profiles (Gong et al., 2003; Krissian et al., 2000; Rohr and Wörz, 2006;

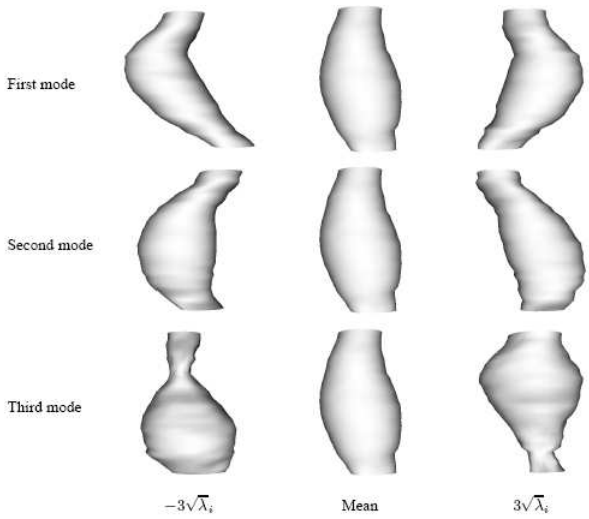
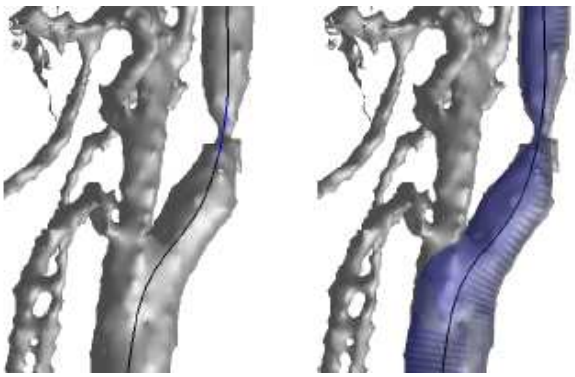


Figure 3.1: Active shape model from (de Bruijne et al., 2003a,b). Modes of variation of the model learned from a database of abdominal aortic aneurysms. Figure based on material from (de Bruijne et al., 2003a).



(*a*) (*b*)

Figure 3.2: B-spline centerline and surface from (Frangi et al., 2000). (a) Centerline B-spline after optimization. (b) 3D B-spline surface deformed with respect to to the centerline and underlying data. Application to carotid arteries in contrast enhanced MRA. Figure based on material from (Frangi et al., 2000).

Shikata et al., 2004), for which saliency criteria (Steger, 1996) and estimation bounds (Rohr and Wörz, 2006) were analytically derived. In practice, vessels (especially large ones) often do not have rounded but rather plateau-like profiles, which are not adequately described by the aforementioned models. The simplest model accounting for that observation is the *bar-like profile* (Boldak et al., 2003; Koller et al., 1995; Steger, 1998), which corresponds to a local cylinder with homogeneous intensity. Of particular interest are the analytical developments in (Steger, 1998) and the use of an *asymmetric* profile for heterogeneous surroundings, with analytical derivation of the corresponding bias correction.

Bar-like profiles do not account for phenomena such as spatial blur and partial volume effect. A first approach to integrate such effects in a bar-like model is the *3D Phi Line* from (Rohr and Wörz, 2006), based on the Phi Gaussian error function. Another popular model is the *bar-like convolved* profile, where the cross-section radial appearance is defined as a plateau convolved with a Gaussian kernel (Krissian et al., 2000; Wörz and Rohr, 2004, 2007, 2008) (see Fig. 3.3). Estimation bounds are provided in (Wörz and Rohr, 2008). In such models, spatial blur is considered as acquisition-dependent. As previously stated in Sec. 3.1.1, such effects are manifestations of the actual point spread function (PSF) of the imaging device. In this regard, the bar-like convolved model can be seen as a PSF appearance model combined with a 3D cylinder geometric model. Estimation of such models can thus be considered as an inverse, deconvolution problem aiming at estimating the vessel's location, direction and radius given the geometric and PSF models (see also Sec. 3.2.2).

Finally, the work of (Hoogeveen et al., 1998) focuses on the fundamental limits to the accuracy of vessel diameter estimations from MR time-of-flight (TOF) and phase contrast (PC) acquisitions. They base their numerical study on 3D appearance models taking into account acquisition-dependent factors such as flow velocity, radio-frequency pulses, relaxation times and flip angles. These models assume a 3D tubular geometry, with laminar flow along the principal axis. Different intensity profiles are proposed for TOF and PC acquisitions.

Vessels as Ridges Another important body of literature deals with vessels as local ridges of the image hyper-surface (with intensities considered as an extra dimension added to the spatial ones). Such models can be viewed as extensions of Blum's medial representation (Blum, 1967) from binary to gray-level images, aiming at exploiting both edge- and region-based information. Although they generally do not integrate explicit appearance models, ridge models qualify as hybrid models as they correlate spatial and intensity information by specifying the desired mathematical properties of ridge locations. Various ridge definitions have been proposed (Eberly et al., 1994), among which *height ridges*, based on the idea of estimating intensity maxima in special directions. 3D, multiscale extensions of height ridges were exploited in several successful vascular-dedicated techniques (Aylward et al., 1996; Aylward and Bullitt, 2002; Fridman et al., 2003). We refer the readers to (Eberly et al., 1994) for more details on the mathematical framework of the large family of ridge models.

Template-Based Shape Spaces Frangi et al. (1998) introduce a shape space based on second order intensity variations (Hessian matrix analysis), further refined in (Lin, 2003). Shape prototypes are defined through relative variations of Hessian eigenvalues. Vessels are matched to bright, elongated ellipsoids (see Fig. 3.4). Hessian-based features to detect such shapes in the image are discussed in Sec. 3.2.2. A similar shape space is defined in (Hernàndez Hoyos, 2002; Hernàndez Hoyos et al., 2005) based on inertia moments.

Related to generalized cylinders (Sec. 3.1.2), *superellipsoids* (Tyrrell et al., 2007) locally describe the vessel geometry in combination with a constant intensity appearance model (see Fig. 3.5). Instead of being embedded in an image tensor (Hessian or inertia matrix), the parameterization (3D pose and appearance) is made explicit. Similarly, the linear templates from (Friman et al., 2008a), based on a Gaussian line appearance model, and the linear contrast operators from (Lacoste et al., 2005, 2006) are constructed from explicitly parameterized 3D cylinders.

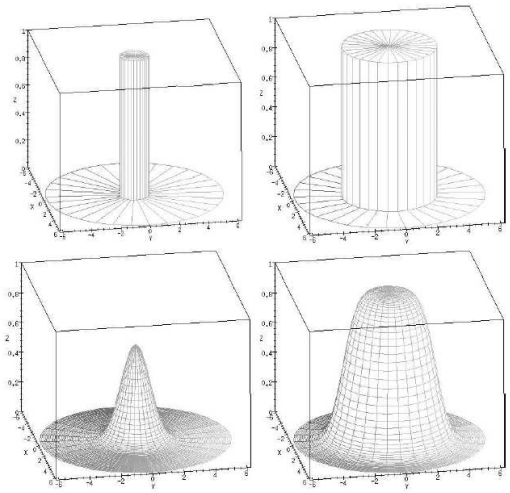


Figure 3.3: Bar-like convolved model from (Krissian et al., 2000; Wörz and Rohr, 2004). Top row: bar-like models of radiuses 1 and 3 voxels. Bottom row: corresponding models after convolution with a Gaussian kernel of standard deviation 0.7. Illustration based on material from (Krissian et al., 2000).

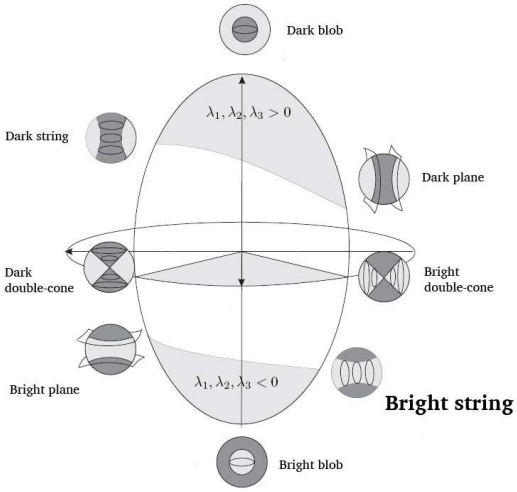


Figure 3.4: Second order ellipsoid shape space from (Frangi et al., 1998; Lin, 2003). Shape prototypes are defined according to Hessian matrix eigenvalues variations (λ_i). Vessels are typically matched to the *bright string* prototype. Illustration based on material from (Lin, 2003).

3.1.4 Models of Bifurcations and Anomalies

Besides regular vessel segments, bifurcations and medically critical anomalies such as calcifications, aneurysms, stents and stenoses can also benefit from dedicated modeling. The scattered apparition of bifurcations and anomalies in angiographies is generally hard to predict, so precise modeling can considerably help their robust detection.

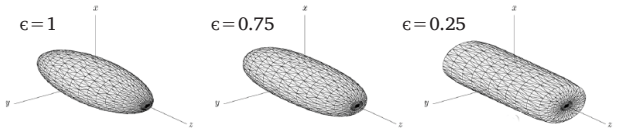


Figure 3.5: Superellipsoids from (Tyrrell et al., 2007). Examples of parameterization of superellipsoids, from an ellipsoid to cylindroids. Illustration based on material from (Tyrrell et al., 2007).

Bifurcations Theoretical geometric bifurcation models were first developed in (Murray, 1926), including relationships between branching angles and vessel widths based on physiological *optimality* conditions (see Fig. 3.6). This model has been evaluated experimentally in various publications such as (Fanucci et al., 1988; Zamir and Brown, 1982, 1983) and further extended in (Oka and Nakai, 1987). Accurate bifurcation models are also proposed in (Antiga and Steinman, 2004; Flaaris et al., 2004), but are not directly suited for segmentation purposes as they require a pre-extracted surface mesh and centerline for their optimization. Simpler geometric models can be found in (Agam and Wu, 2005; Agam et al., 2005), where possible bifurcations are categorized as Y or X junctions (with three and four principal directions, respectively). In (Hernàndez Hoyos, 2002; Hernàndez Hoyos et al., 2005), bifurcations are defined *a contrario* as deviations from cylindrical segments in terms of moments of inertia. In practice, most works do not model bifurcations explicitly. Numerous techniques focus solely on single branch segmentation, while other employ dedicated detection schemes, based notably on topology analysis.

Calcifications and Stents In contrast-enhanced CTA data, calcifications and stents appear as hyper intense structures and are classically handled with specific appearance models, typically resulting in the definition of higher-bound thresholds (Boldak et al., 2003; Wesarg and Firle, 2004).

Stenoses and Aneurysms Stenoses are generally modeled implicitly as local and sudden radius decrease (Frangi et al., 2000; Hernàndez Hoyos et al., 2000; Sato et al., 1998a; Sun et al., 1995; Verdonck et al., 1996). Among explicit stenosis models useful for segmentation, one can mention the double cone model (Frangi et al., 1998; Hernàndez Hoyos, 2002; Hernàndez Hoyos et al., 2005; Lin, 2003) used in (Lin, 2003) to derive a specialized detector based on Hessian analysis.

Aneurysms, due to their high shape variability, are generally not explicitly modeled. A counter example is the work from (de Bruijne et al., 2003b), where a model of abdominal aortic aneurysms is learned from a database (see Fig. 3.1). More generally, special effort is often put into making methods robust to aneurysms (Lin, 2003; Wink et al., 2000a), so that they can be detected *a posteriori* as local deviations from the vessel model. Experiments from the literature suggest local radius and volume increase as possible indicators of most prominent aneurysms. Refined morphological criteria are employed in (Kawata et al., 1995).

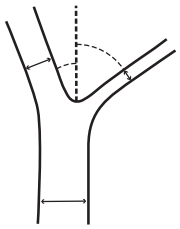


Figure 3.6: Geometric model for bifurcations. Geometrical models of Murray (Murray, 1926) and Oka and Nakai (Oka and Nakai, 1987) specify relations between vessel widths and angles based on physiological considerations. Illustration based on material from (Hernàndez Hoyos, 2002; Hernàndez Hoyos et al., 2005).

3.1.5 Problem-Specific Models

Application-, organ-specific information may also be incorporated to ease the segmentation task. We already mentioned the work of (de Bruijne et al., 2003b) on abdominal aortic aneurysms (see Fig. 3.1) as an example of application-specific learned model. In (Florin et al., 2004), authors use the heart surface to constrain a ray-casting/profile analysis to likely locations of coronaries in cardiac CT angiographic data (see Fig. 3.12). In (Shikata et al., 2004), a lung mask is used for pulmonary artery segmentation. In the relatively rare applicative contexts where reliable statistical data is available, atlases can be employed to guide vascular segmentation, as done in (Passat et al., 2006) for brain vasculature.

3.1.6 General Considerations on Vessel Models

Numerous factors influence practical model choices. Prior knowledge on the appearance of the vessels of interest is often specific to the target modality. On the other hand, the adequacy of a geometric model depends on the application (size, regularity and complexity of the target vessels), but also on the eventual clinical goals. Simple centerline models are generally sufficient for visual screening tasks, but more complete, accurate representations such as generalized cylinders are required for the quantification of vascular pathologies.

A single segmentation algorithm may naturally rely on several complementary models. For instance, a hybrid model can be used for local geometry and appearance, with generalized cylinders controlling the geometric coherence of the result at a higher level. Models can also be combined in coarse-to-fine, multi-layer designs. A simple centerline model can guide a coarse, but fast and robust delineation of the vessels of interest. Such a rough delineation can then be used to efficiently initialize and constrain a refined segmentation process, governed by a more complex model. Such an approach was for instance exploited in (Frangi et al., 1999a, 2000) and benefit to the overall robustness and computational efficiency of the segmentation scheme.

3.2 Vascular-Dedicated Features

Hereafter, we review existing vessel-specific features, *i.e.*, the practical detectors and filters used to evaluate vascular models on the data. For each technique, we describe which

kind of information is exploited and how it is extracted. Besides their theoretical background, we discuss their robustness, accuracy, discriminative power and computational efficiency.

Vessel-dedicated features are commonly designed by combining and adapting simple, well-spread image processing tools, from basic image intensities to first and second order edge-based detectors (*e.g.* gradients, Canny detector, Laplacian zero-crossings). The reader may refer to general and medical image processing reviews and handbooks (Ayache, 1995; Duncan and Ayache, 2000; McInerney and Terzopoulos, 1996; Sonka et al., 1996; Sonka and Fitzpatrick, 2000) for extensive overviews of these classical techniques.

Features of interest incorporate vessel-specific prior knowledge (typically derived from an underlying model) in their designs, *e.g.* assumptions on the elongation and the direction, on the scale and circularity of vascular cross-sections. Scale and directionality are two aspects essential to *multiscale line filters*, from which numerous vessel features are derived. We refer the reader to (Lindeberg, 1994; ter Haar Romeny, 2003) for comprehensive overviews of line detection in scale-space frameworks.

3.2.1 Local Isotropic Features

A few vessel-dedicated features from the literature solely focus on estimating the location and/or scale of target vessels. We refer to such features as *isotropic*, as they do not exploit assumptions on the directionality of the vessels.

In contrast-enhanced modalities, the simplest assumption one can exploit is that vessels are brighter than their immediate surroundings. In (Szymczak et al., 2005), authors rely on local, scale-persistent *robust intensity maxima* to select a set of candidate points, expected to lie mostly inside the vessels. Similarly, local intensity maxima on 2D slices are selected as seed points for the tracking scheme of (Lee et al., 2007).

In (Aylward et al., 1996), authors implement a 3D, multiscale ridge detector (see Sec. 3.1.3). This *medialness* operator is based on multiscale Laplacians of Gaussians, assuming that vascular intensities are locally maximum. Extensions of this work (Aylward and Bullitt, 2002; Fridman et al., 2003; Furst, 1999), explicitly using vessel directionality, are presented in Sec. 3.2.3.

Another noticeable technique in this category is the multiscale spherical flux measure first introduced in Vasilevskiy and Siddiqi (2002). This feature approximates the divergence of image gradients by measuring the gradient flux through the boundaries of multiscale spheres. Its response is maximal at central vessel locations and corresponding scale. It was proved to be able to detect even low contrast and narrow vessels. An efficient Fourier-based implementation of the spherical flux feature was recently proposed in Law and Chung (2009).

3.2.2 3D Local Geometry Features

Besides hyper-intensity, a core characteristic of vessels is their specific local geometry. Techniques have thus been proposed to locally analyze image data in 3D, with the aim of distinguishing hyper-intense, tubular-like patterns.

Derivative Features A popular approach in the detection of vascular patterns is the use of second-order derivative information (principal curvatures of image intensities) to characterize the local image geometry. The core assumption of second-order derivative

methods is that vessels can be distinguished by a locally prominent low curvature orientation (the vessel direction) and a plane of high intensity curvature (the cross-sectional plane). The Hessian matrix is the most common tool to capture such information. To the authors' knowledge, it was first used for vascular segmentation in (Koller et al., 1995) to estimate vessel orientations thanks to Hessian eigenvectors. Following this seminal idea, a fair amount of work was dedicated to multiscale Hessian-based filters for vessel enhancement (Frangi et al., 1998; Li et al., 2003; Lin, 2003; Lorenz et al., 1997; Sato et al., 1998b; Shikata et al., 2004). These filters rely on Hessian eigenvalues to discriminate between plane-, blob- and tubular-like structures, according to prototype-based shape models (see Sec. 3.1.3 and Fig. 3.4). The *vesselness* filter from (Frangi et al., 1998) (see Fig. 3.7) has been extensively used in practice, owing to its intuitive geometric formulation. The *Weingarten* matrix is a less popular alternative to the Hessian matrix (Armande et al., 1996; Prinet et al., 1996).

Hessian- and Weingarten-based filters may suffer from sensitivity to local deformations. Aneurysms, stenoses and bifurcations are generally reported to produce false negative responses as such situations sensibly deviate from the assumption of a single strong orientation. Noise and nearby non-vascular structures may also perturb the response of such filters.

These techniques commonly rely on multiscale frameworks to cope with different vessel sizes. Hessian and Weingarten matrices are classically computed in a Gaussian linear scale-space, through convolutions by Gaussian derivatives of different standard deviations (Lindeberg, 1994). Scale selection is performed by keeping the maximum response over multiple scales. The extraction of large vessels requires large standard deviations which induce substantial spatial blur. The response of these filters can thus be perturbed by the presence of other hyper-intense structures in the immediate surroundings of the target vessel. Finally, depending on the range of selected scales, these filters may yield substantial computational costs.

The recent work of (Bauer and Bischof, 2008c,b) is an interesting alternative to multiscale Hessian-based filters. The Hessian matrix can be viewed as the Jacobian of the gradient vector field. Instead of relying on the Gaussian scale-space gradient vector field, Bauer and Bischof (2008c) propose to use an other vector field, obtained from *Gradient Vector Flow* (GVF) diffusion (Xu and Prince, 1998). The GVF scheme regularizes and anisotropically diffuses gradients from the object boundaries. This vector flow vanishes at medial locations, *i.e.*, on vessel centerlines. Bauer and Bischof (2008c) apply the vesselness criteria from (Frangi et al., 1998) on the local Jacobian of the GVF field to derive their enhancement filter. This approach overcomes most of the shortcomings of Gaussian scale-space techniques. It is mono-scale and less sensitive to perturbations from structures adjacent to the target vessels. To alleviate the high computational cost of GVF diffusion, Bauer and Bischof (2008c) use a GPU implementation.

Instead of analyzing second-order derivatives of the image, one can alternatively exploit the spatial covariance of image gradients vectors. A first subgroup of such methods focuses on the local distribution of gradient vectors through the analysis of the *structure tensor* matrix (covariance matrix of gradient vectors) (Weickert, 1999) to exhibit clusters of gradient vectors. Works from (Agam and Wu, 2005) and (Agam et al., 2005) share the ability of detecting potentially more than one principal direction. They can thus discriminate between simple vascular segments, nodules and bifurcations. In (Agam and Wu, 2005), estimations are performed through Expectation-Maximization of a probabilistic mixture model, while iterative component analysis is preferred in (Agam et al., 2005). In

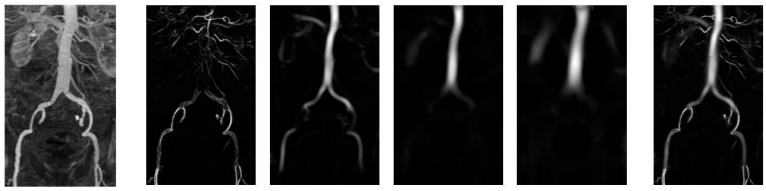


Figure 3.7: Hessian-based multiscale filter from (Frangi et al., 1998). Application to the enhancement of aortoiliac MRA. From left to right: original image, four increasing scale response images, combined in the final enhanced image. Illustration based on material from (Frangi et al., 1998).

terms of scale adaptation, (Agam and Wu, 2005; Agam et al., 2005) implement heuristic rules to select the size of the window of interest. In (Agam et al., 2005), authors favorably compare their methods to Hessian-based filters (Frangi et al., 1998; Li et al., 2003; Shikata et al., 2004), reporting smaller mean square localization errors and a better behavior at bifurcations.

Another recent approach exploiting the distribution of gradient vectors is the *optimally oriented flux* from (Law and Chung, 2008). Similarly to the spherical flux Vasilevskiy and Siddiqi (2002); Law and Chung (2009) discussed in Sec. 3.2.1, the optimally oriented flux relies on the measure of gradient flux through the boundary of local spheres. The key difference lies in the exploitation of the vessel directionality through the estimation of an optimal gradient projection axis. Law and Chung (2008) develop a Fourier-based implementation and discuss the advantages of optimally oriented flux over Hessian-based measures, including more accurate and stable responses and higher robustness to disturbances from adjacent structures.

Closely related, several oriented features use first order derivative information along 2D cross-sectional circular patterns (Fridman et al., 2003; Fridman, 2004; Gulsun and Tek, 2008b; Krissian et al., 2000; Wink et al., 2000a). These features are explicitly oriented, *i.e.*, the direction to be tested is an input parameter. Such cross-sectional features are discussed in Sec. 3.2.3.

Integrative Features As an alternative to derivative methods, some authors proposed to use integrative measures such as second-order *inertia moments* of image intensities (Boldak et al., 2003; Hernàndez Hoyos et al., 2000; Hernàndez Hoyos, 2002; Hernàndez Hoyos et al., 2005, 2006; Larralde et al., 2003; Reuze et al., 1993). These filters are theoretically less sensitive to noise while providing similar information on local geometry. Vessel detection filters can be designed similarly to Hessian-based ones, with ratios of the eigenvalues of the inertia matrix (Hernàndez Hoyos et al., 2000; Hernàndez Hoyos, 2002; Hernàndez Hoyos et al., 2005, 2006). Contrarily to second order derivative methods, integrative techniques do not enjoy a well-defined scale-space theoretical framework. Most works use a fixed, single scale. Heuristics to adapt its value are proposed in (Hernàndez Hoyos et al., 2000; Hernàndez Hoyos, 2002; Hernàndez Hoyos et al., 2005, 2006).

In the binary case, the inertia matrix reduces to the spatial covariance matrix of the binary shape. In (Lorenz et al., 2003), such a covariance matrix is used to iteratively check pre-extracted vessel segments and heuristically adapt segmentation parameters. The scale-independent *elongation* criterion of (Wilkinson and Westenberg, 2001) also uses spatial

covariance for vessel enhancement. This criterion is applied to the connected components given by the gray levels of the image, in the context of *morphological connected filters* (Serra and Salembier, 1993).

Local Fit of 3D Hybrid Models In Sec. 3.1.3, we described how radial appearance profiles can be used to model intensity variations locally, in tubular regions of interest. If such a hybrid model can be expressed (or approximated) in a closed form, partial derivatives with respect to the parameters can be derived analytically. Then, the hybrid model can be directly fitted to the local data, typically with non-linear least square optimization such as *Levenberg-Marquardt* schemes. This method was successfully applied for Gaussian profiles in (Friman et al., 2008a) and bar-like convolved models in (La Cruz et al., 2004; Wörz and Rohr, 2004, 2007, 2008). Goodness of fit at convergence can be used as a measure to discriminate between vascular and non-vascular areas. Measures of fit significance, evaluated with *Student t-tests*, can additionally be derived Friman et al. (2008a).

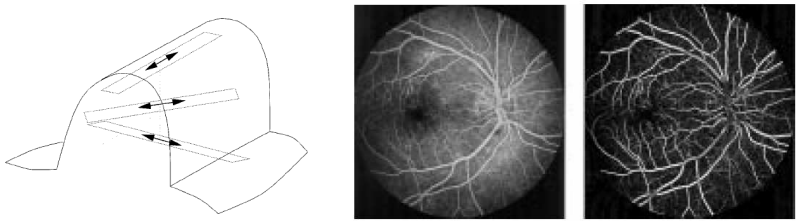
Although not relying on an explicit hybrid model, the framework of (Lacoste et al., 2005, 2006) also uses linearly oriented elements as local contrast operators. *Student t-tests* quantify the significance of contrast values. No local optimization is performed, as the segmentation process is globally driven by a *Markov marked point process* (see Sec. 3.3.5). In (Tyrrell et al., 2007), local *superellipsoids* (see Fig. 3.5) rely on an estimation of the local contrast using a piecewise constant appearance model (constant object on a constant background). Local optimization is performed through log-likelihood maximization. Significance is measured through generalized likelihood ratio tests, in the context of robust M-estimators. Authors emphasize the robustness of this estimation scheme to nearby structures such as adjacent vessels.

***A Contrario* Features** The previous features rely on the general principle of maximizing the fit of a vascular model. Another interesting approach is to detect vessels *a contrario*. In Sec. 3.1.2, we already mentioned the *ball measure* from (Nain et al., 2004) and the non-local measures from (Rochery et al., 2005a, 2006). These purely geometric features are used to guide active contour optimizations by penalizing deviations from tubular-like shapes (see Sec. 3.3.3 and Fig. 3.18 and 3.19).

Another feature in this category is the top-hat filter. This morphological operator detects local intensity peaks through their *suppression* by morphological opening. Its general formulation can be tailored to exploit the thinness and elongation of vascular structures by using linear oriented patterns as structuring elements. When oriented perpendicularly to a vessel in the image, such oriented patterns will cause its suppression by opening, hence its detection by the top-hat filter. The directional top-hat sum technique from (Zana and Klein, 2001) is illustrated in Fig. 3.8.

3.2.3 2D Cross-Sectional Features

Some authors focused their analysis on the cross-section geometry, testing hypotheses such as its compactness and regularity. Cross-sectional features can be viewed as special cases of 3D local geometry features from Sec. 3.2.2. They have the advantage of downscaling the analysis from three to two dimensions. The orientation to be tested are considered as an input parameter. Directional information is commonly estimated from features such as Hessian eigenvalues (see Sec. 3.2.2) or given by the extraction scheme (see Sec. 3.3). A



(a) (b)

Figure 3.8: Oriented top-hat from (Zana and Klein, 2001). (a) Opening with linear structuring elements. The 2D vessel is illustrated as a 3D crest line (intensity elevation map). Three structuring elements are given. The structuring element perpendicular to the vessel will give the best top-hat response, as it can not be contained in the vessel and will remove it from the opening. The two other directional openings will leave the vessel nearly unaltered. (b) Left: original image (2D retinography). Right: sum of oriented top-hats. Illustrations based on material from (Zana and Klein, 2001).

single orientation estimation may be sensitive to noise and local perturbations. A robust but potentially costly alternative is to test exhaustively a discrete set of possible directions.

In (Behrens et al., 2001), a randomized 2D ellipse Hough transform is used to detect 2D cross-sections on edge maps. While computationally efficient, such a technique does not provide very accurate localization and heavily relies on the underlying edge detector. Following similar ideas, approximate circle fitting, based on maxima selection of a multiscale medial filter, is proposed in (Rueckert et al., 1997).

In (Florin et al., 2005), the so-called *ribbon* measure compares the mean intensity inside a 2D ellipse with its immediate surroundings. This contrast-based feature reaches local maxima when the ellipse is well aligned with the vessel cross-section. Parameters are optimized in a *particle filter* framework (see Sec. 3.3.5).

Cross-Sectional Circular Medialness Measures In the context of ridge tracking (see Sec. 3.1.3), the work from (Aylward and Bullitt, 2002) extends their original *medialness* feature (Aylward et al., 1996) by exploiting the vessel's directionality, through a 2D convolution kernel measuring the contrast on circular cross-sections. Direction is estimated through the Hessian eigenvectors and a new scale optimization scheme is proposed.

The work of (Koller et al., 1995) on steerable derivative filters introduced the seminal ideas shared by many cross-sectional medialness features. Koller et al. (1995) base their initial 2D analysis on the optimal line detector from (Canny, 1986) convolved with a bar of a variable width. They rely on two shifted, oriented Gaussian derivative kernels to detect parallel vessel walls. A multiscale scheme handles varying vessel widths. Authors discuss the theoretical properties of their filters, such as the response to different models of vessel profiles (Sec. 3.1.3). Of particular interest, Koller et al. (1995) propose a non-linear combination (through a *min* operator) of the shifted kernels to reduce step-edge false responses. The extension to 3D was the first vascular filter, to our knowledge, to use the Hessian matrix to estimate the local orientation. Only two directions are tested in the cross-sectional plane, whose responses are again combined with a *min* operator.

Very similar ideas are developed in (Poli and Valli, 1996), in 2D, with a particular focus on kernel separability for computational efficiency. As noted by the authors, these linear

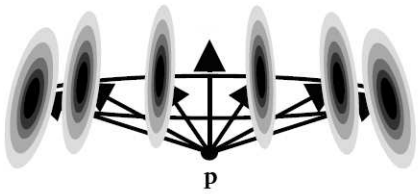


Figure 3.9: Core medial atom from (Fridman et al., 2003). The length of equi-angular spokes corresponds to the radius of the vessel being tested. Weighting functions (directional Gaussian derivatives), at the end of each spoke, are summed to obtain the filter response. Illustration based on material from (Fridman et al., 2003).

filters are subject to false positive responses at step edges. A basic intensity comparison between "left" and "right" positions is proposed to alleviate that shortcoming. Another closely related approach is the so-called *ellipsoidal scale-space* filter from (Suri et al., 2002c).

Inspired by (Aylward and Bullitt, 2002; Koller et al., 1995) and (Furst, 1999), the work from (Fridman et al., 2003; Fridman, 2004) develops a formulation of the *core* framework (Pizer et al., 1998) dedicated to vascular structures. A medial atom response is obtained by convolving the image with directional Gaussian derivative kernels located at the end of equi-angular radial spokes (see Fig. 3.9). Given a location, an orientation and a radius, the core feature measures the fit of the image edges with a circular cross-section model. Bifurcations are handled thanks to an additional edge detector coupled with specific patterns of the core response. Similarly, the adaptive filter from (Krissian et al., 2000) uses a sum of oriented gradient responses along a circle. In (Krissian et al., 2000), special attention is given to the analytical and experimental study of the filter with respect to scale estimation and response to different appearance radial profiles. Orientation is classically given by Hessian eigenvectors.

Ray-Casting Features The idea of ray-casting has received much interest for the characterization of 2D vessel cross-sections (Gulsun and Tek, 2008b; Niessen et al., 1999; Tek et al., 2001; Verdonck et al., 1995; Wesarg and Firle, 2004; Wink et al., 1998, 2000a). Given a point p in the vessel and a direction $\overrightarrow{d}$ defining the cross-sectional plane, ray-casting techniques aim at detecting the vessel wall along rays thrown equi-angularly from p. Detected contour points can then be used to evaluate the centerness of a point inside a vessel. Existing ray-casting features are closely related to the medialness measures from the previous paragraph. Instead of relying on a circular or elliptic model, they only assume a compact, regular cross-sectional contour.

As depicted in Fig. 3.10, ray-casting reduces the detection problem to a 1D analysis along the rays, yielding in general high levels of computational performance. Intensity (Wesarg and Firle, 2004) or gradient thresholding (Tek et al., 2001; Wink et al., 2000a) is classically used to select candidate contour points. Robustness is ensured by casting and averaging numerous rays. Additionally, constraints on the spatial coherence of neighboring contour points can be enforced through simple heuristics, or, more globally, thanks to an optimization by dynamic programming (Verdonck et al., 1995). The center of mass of candidate contour points is often employed to measure the centerness of a location

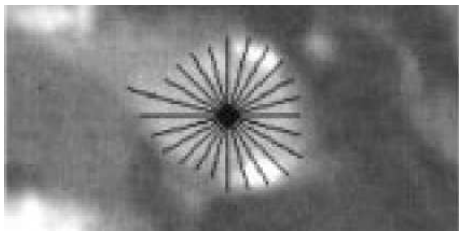


Figure 3.10: Ray-casting technique from (Wink et al., 2000a). From a point in the vessel, rays are cast in a 2D cross-section and gradient maxima are selected as contour points. Note the robustness to calcifications. Application to CT angiography. Illustration based on material from (Wink et al., 2000a).

(Niessen et al., 1999; Verdonck et al., 1995; Wesarg and Firle, 2004; Wink et al., 1998, 2000a). A notable alternative is the use of *mean-shift* filtering in (Tek et al., 2001).

In (Gulsun and Tek, 2008b), ray-casting is used to design a multiscale medialness operator, similarly to other filters described at the beginning of this section. Responses along rays are summed in a distance-wise manner, following a circular cross-section model. A local normalization scheme is employed to convey contrast invariance. This new medialness measure makes use of the sign of the gradient along a ray to lessen the influence of nearby hyper-intense structures. Gradient polarity along rays was also exploited in (Wesarg and Firle, 2004; Wink et al., 2000a) to handle hyper-intense anomalies such as calcifications.

Accurate Extraction of 2D Contours Techniques performing an accurate segmentation of 2D cross-sectional contours can also be considered as special cases of cross-section features. Similarly to most ray-based techniques, such methods do not rely on fixed geometric models. Instead, their optimization schemes integrate soft compactness and regularity constraints on the cross-sectional contours. Classical techniques in that category include 2D parametric active contours, as used in (Hernàndez Hoyos, 2002; Hernàndez Hoyos et al., 2005; Li and Ourselin, 2003) (see Fig. 3.28). In (Rueckert et al., 1997), active contours are coupled with stochastic relaxation to overcome local minima issues. One can also highlight the family of circular dynamic-programming methods, such as the 2D *star pattern* from (Geiger et al., 1995), used in (O'Donnell et al., 1998; Verdonck et al., 1995), the *circular shortest path* from (Appleton and Sun, 2003) used in (Florin et al., 2006) and the *mean graph cycle optimization* used in (Gulsun and Tek, 2006, 2008b,a). These techniques benefit from optimal and relatively efficient optimization schemes.

3.2.4 Features for Bifurcations and Anomalies

Bifurcations Bifurcations can be detected by the emergence of several principal directions, as exploited by the clustering method in (Agam et al., 2005). In (Hernàndez Hoyos, 2002; Hernàndez Hoyos et al., 2005), branching point candidates are selected through a inertia moment-based heuristic. These ideas are further refined in (Carrillo et al., 2007), with the analysis of connected components on a local sphere surface through a moment-based clustering algorithm. This notion of binary connected components was also exploited in (Flasque et al., 2001), with a local parallelepiped ROI encompassing the extracted segment. In (Fridman et al., 2003), authors use responses of a *corner detector* (Blom, 1991) combined with a bifurcation-specific adaptation of their *core* filter. If the features provide

a width estimate, local increase of this value can be used as a clue for the detection of bifurcations (Aylward and Bullitt, 2002). Most other branch handling schemes are based on *a posteriori* topological analysis of segmented results and are dependent on the extraction method (Sec. 3.3).

Stenoses and Aneurysms Lin (2003) propose a stenosis detector based on Hessian analysis. Stenosis detection is more generally performed *a posteriori* by quantifying radius decrease (Frangi et al., 2000; Hernàndez Hoyos et al., 2000; Sato et al., 1998a; Sun et al., 1995; Verdonck et al., 1996), by blood flow simulation (Deschamps et al., 2004) or by 3D curvature analysis (Kawata et al., 1995, 1996).

Similarly to stenoses, aneurysms can be detected *a posteriori* as local radius increases, or by 3D curvature analysis (Kawata et al., 1995, 1996). Although none has been proposed to the authors' knowledge, one could probably design Hessian-based aneurysm detectors, as hinted in (Sato et al., 1998b). Other techniques dedicated to aneurysms include learned shape models (de Bruijne et al., 2003b,a) and volume-based estimation (Hernandez et al., 2003).

Calcifications Calcifications are usually handled simply as hyper-intense areas inside the vessel. In (Wesarg and Firle, 2004), calcifications are separated from the vessel lumen by dual thresholding. In (Wink et al., 2000a), calcifications are detected as high positive gradient along concentric rays in 2D cross-sections.

3.2.5 General Considerations on Vessel Features

Existing vascular features differ naturally in terms of accuracy, robustness and computational efficiency. Another important consideration to take into account is the amount of information that a given feature extracts from the image. For instance, popular Hessian-based detectors require only test locations to be specified and provide both direction and scale estimations. However, these low-level cues are potentially corrupted by local image perturbations. 2D cross-sectional features generally provide robust scale estimations, but require both location and direction as inputs. Input parameters can be viewed as defining the search space to be explored by the overall extraction process. Specialized features, with several parameters, yield more complex optimization schemes. On the other hand, the extraction process can, in general, incorporate higher-level, model-based information to compensate the unreliability of low-level, feature-based estimations. The balance between low-level, high-level, data-based and model-based information is a critical and intricate design issue.

3.3 Extraction Schemes

In this section, we review the optimization schemes proposed in the literature to effectively perform vessel segmentation. We discuss possible algorithmic designs, relying on models from Sec. 3.1 and exploiting features from Sec. 3.2. First, we briefly review some *pre-processing* and *pre-segmentation* approaches to simplify image content, improve image quality and obtain an approximate localization of the vessels (Sec. 3.3.1). Such preliminary results may serve as inputs for vessel-dedicated extraction schemes, which we separate in three general classes:

– region-growing algorithms (Sec. 3.3.2);

– active-contours (Sec. 3.3.3);
– centerline-based approaches (Sec. 3.3.4).

Additionally, we highlight two promising stochastic techniques recently applied to vascular segmentation: particle filtering and Markov marked point processes (Sec. 3.3.5). We finally discuss various *post-processing* techniques to refine segmentation results in Sec. 3.3.6. The reader may refer to Fig. 3.11 as a guide for this section's layout.

3.3.1 Pre-Processing and Pre-Segmentation

Pre-Processing To simplify the tremendous amount of information embedded in 3D angiograms, one can first consider generic techniques such as down-sampling and quantization (Tschirren et al., 2005). One can also improve the image quality through filtering. Numerous features from Sec. 3.2.2 were originally proposed as vessel enhancement filters (see Fig. 3.7). This includes derivative filters from (Koller et al., 1995; Poli and Valli, 1996), structure tensor analysis (Agam and Wu, 2005; Agam et al., 2005), mathematical morphology filters (Wilkinson and Westenberg, 2001; Zana and Klein, 2001), popular Hessian-based filters (Frangi et al., 1998; Krissian et al., 2000; Li et al., 2003; Lin, 2003; Lorenz et al., 1997; Sato et al., 1998b; Shikata et al., 2004) and flux-based filters (Law and Chung, 2008, 2009). Another filtering technique not previously mentioned is the vessel enhancement method from (Westin et al., 2001), which relies on a direction estimation through frequency-based *quadrature* filters to perform locally adaptive low-pass and high-pass filtering.

More generally, a fair amount of research effort can be found in vessel-dedicated *anisotropic diffusion* schemes (Gerig et al., 1992; Krissian et al., 1997; Manniesing et al., 2006; Meijering et al., 2001; Weickert, 1999) for angiography enhancement. These regularization schemes cope with the challenge of reducing noise levels while preserving thin and weakly contrasted vessels. Comprehensive overviews and evaluations of state-of-the-art approaches are proposed in (Meijering et al., 2001) and (Suri et al., 2002b). Diffusion is guided by the local anisotropy of the data, classically derived from the Hessian matrix (Krissian et al., 1997; Manniesing et al., 2006) or structure tensors (Weickert, 1999). Vectorial gradient *flux* was also used for anisotropic diffusion and was demonstrated to be successful at preserving thin, low-contrast vessels (Krissian, 2002). Finally, works on vascular-dedicated filtering from (Orkisz et al., 1997, 2000; Tankyevych et al., 2008) should also be highlighted as employing non-linear, anisotropic filtering based on local estimation of main orientations.

Pre-Segmentation Most of the extraction algorithms of the next sections require some rough, prior localization of the target vessels. Regions of interest (ROIs) can first be selected given prior anatomical knowledge. Examples include the use of a mask of the heart for coronary segmentation in (Florin et al., 2004) (Fig. 3.12), a mask of the lungs for pulmonary vessels in (Shikata et al., 2004) and a probabilistic atlas for cerebral vasculature in (Passat et al., 2006). ROIs can also be obtained by thresholding a *potential map*. Image intensities were directly exploited in (Wu et al., 2004). A more advanced approach is to derive mixture probability maps from an appearance model (Sec. 3.1.1), estimated by Expectation Maximization (Chung et al., 2004; Gan et al., 2004; Hassouna et al., 2003; Wilson and Noble, 1999; Yang et al., 2004) (see Fig. 3.13), or from vessel-enhanced images (Agam and Wu, 2005; Agam et al., 2005; Shikata et al., 2004).

Schemes have also been designed to select only a sparse set of candidate points. Robust

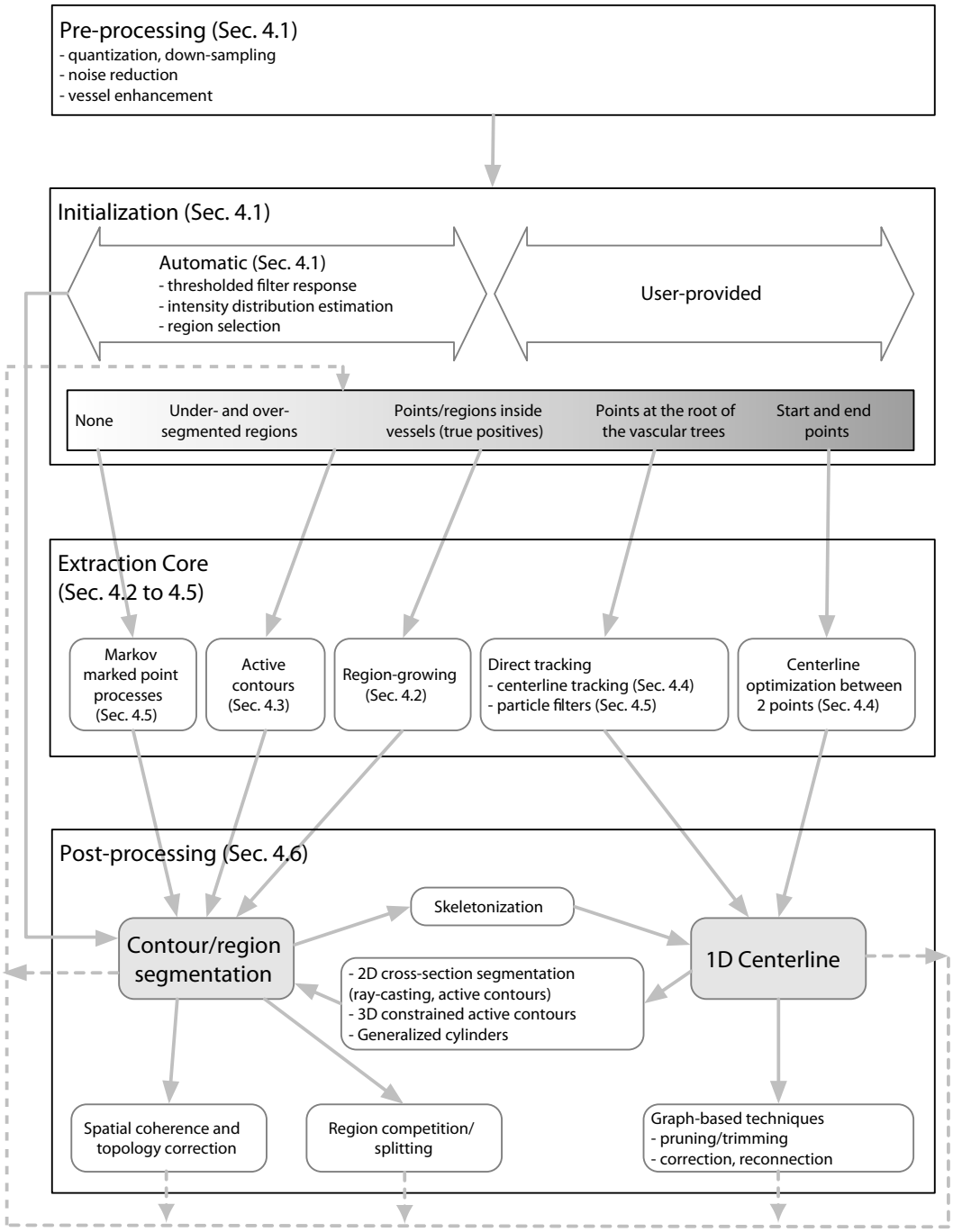


Figure 3.11: Extraction schemes: recapitulative diagram.
The diagram links the different subsections of Sec. 3.3 by exhibiting classical combinations of pre-treatments, initializations, extraction core techniques and possible post-processing. Core extraction techniques (Sec. 3.3.2 to 3.3.5) are displayed in increasing order of input requirements, from left to right. Complex algorithmic designs can be composed of several processing layers, symbolized by the loops in the diagram. For instance, a centerline obtained by direct tracking can serve as input for accurate surface extraction by active contours.

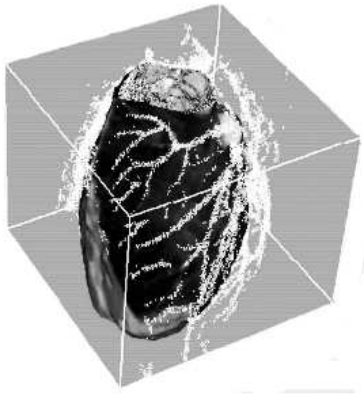


Figure 3.12: Ray-casting candidate selection from (Florin et al., 2004). Rays are cast from the center of the heart and analyzed for intensity peaks. Selected points form clouds likely to be located in cardiac vessels. Illustration based on material from (Florin et al., 2004).

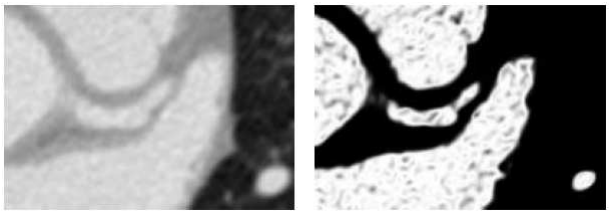


Figure 3.13: EM-based segmentation from (Yang et al., 2004). Left: original image (coronary CTA). Right: maximum *a posteriori* image for blood class, including vessels and heart chambers. Anisotropic smoothing is used to ensure spatial coherence. Illustration based on material from (Yang et al., 2004).

maxima are exploited in (Szymczak et al., 2005) (Fig. 3.29). In (Florin et al., 2004), intensity peaks are detected along 1D rays (Fig. 3.12). Similarly, local maxima are used in (Lee et al., 2007) as seed points for a tracking algorithm (see Sec. 3.3.4).

Most of these pre-segmentation methods are based on pixel-wise processes which may be prone to fragmented outputs, false positive and false negative detections. Correction schemes, ensuring spatial coherence or particular topological properties, are discussed in Sec. 3.3.6.

3.3.2 Region-Growing Approaches

From seed points or regions located inside a vessel, *region-growing* approaches incrementally segment an object by recruiting neighboring voxels based on some inclusion criteria. Some of these schemes require seed points to be located precisely, at the root of the vascular tree for instance. Inputs are then generally provided manually. A key to the popularity of region-growing schemes, besides their simplicity, is certainly their computational efficiency. They fall in the category of greedy algorithms, employ low-level, simple inclusion rules, and explore datasets only sparsely, a critical advantage for large 3D datasets.

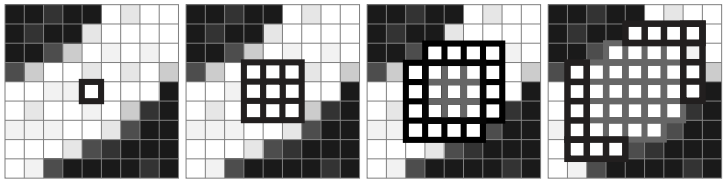


Figure 3.14: Region-growing for vascular segmentation. From left to right: starting from a seed point, a region-growing algorithm successively adds neighboring voxels fulfilling inclusion criteria (*e.g.*, an intensity threshold).

Classical Region-Growing In (Boskamp et al., 2004), an initial segmentation is obtained via region-growing based on a simple intensity threshold (see Fig. 3.14). Such a criterion is sensitive to noise and contrast agent inhomogeneities. As classical region-growing are voxel-wise processes, they are also prone to false negative (holes) and false positive (leakage) issues. In (Metz et al., 2007), growth-limiting criteria are proposed to lessen risks of leakage. Alternatively, dual object and background *competitive region-growing* is proposed in (Yi and Ra, 2003). The morphological dilation-based growing scheme from (Masutani et al., 1998) also aims at reducing such topological issues.

Algorithmic enhancements of classical region-growing implementations can be found for instance in (Wan et al., 2000, 2002b) with *symmetric region-growing*. The *corkscrew* algorithm from (Wesarg and Firle, 2004) can also be considered as a particular variant of region-growing ideas. On a locally thresholded image, linear search paths, cycling standard 3D directions, are followed until the vessel wall is found. Data are explored very sparsely, only along 1D discrete rays. Such a particular search strategy makes an interesting bridge with direct centerline tracking techniques discussed in Sec. 3.3.4.

Wave Propagation Wave propagation techniques, also known as *front propagations*, evolve a well-formed interface inside the vessel (Fig. 3.15). Their advantage over classical region-growing approaches is the spatially coherent propagation they enforce, which eases advanced dynamical and topological analysis. Discrete waves were proposed in (Bruijns, 2001; Cai et al., 2006; Kirbas and Quek, 2002, 2003; Quek and Kirbas, 2001; Tuchschmid et al., 2005; Zahlten, 1995). These can be seen as ordered region-growing schemes, where candidates for inclusion are ordered according to fitness criteria with local schemes to correct the geometry of the growing front (Zahlten, 1995). A more accurate yet still computationally efficient alternative is the *fast-marching* algorithm (Adalsteinsson and Sethian, 1995; Sethian, 1996a; Tsitsiklis, 1995). Initially derived as a special case derivation of general level-set approaches (see Sec. 3.3.3), it can be used as an ordered propagation scheme. Applied on the image grid and given an adequate vessel-dedicated potential, it visits voxel candidates according to an estimation of their geodesic distance from seed points. Another closely related use of fast-marching schemes is the optimization of minimal paths (see Sec. 3.3.4). The various vessel-dedicated fast-marching-based growing methods vary in their definitions of the wave speed function: uniform Euclidean speed and binary inclusion criterion in (Lorenz et al., 2003); weighted geodesic speed in (Sebbe et al., 2003); dynamically adapted speed in (Manniesing and Niessen, 2004). In addition, stabilization schemes, used to correct the orientation of the front in high curvature areas, can be found in (Yan and Zhuang, 2003).

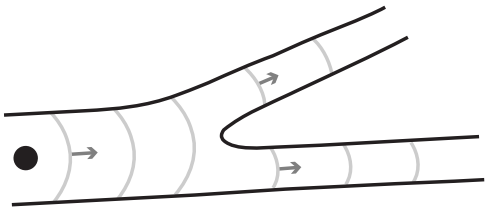


Figure 3.15: Wave propagation for vascular segmentation. From the seed point, a lens-like interface, also known as wave or front, is propagated along the vessel.

On Adaptivity A spatially coherent propagation yields important benefits, such as easier topology analysis. The connectivity of the propagating interface can be exploited to detect bifurcations (Lorenz et al., 2003; Metz et al., 2007; Sebbe et al., 2003; Sekiguchi et al., 2005; Zahlten, 1995) (see Fig. 3.15). Branching handling can be further eased by a local approximation of the vessel centerline, which can be derived from the successive front centers (Lorenz et al., 2003; Tuchschmid et al., 2005), by skeletonization of the local binary mask (Manniesing and Niessen, 2004; Tschirren et al., 2005) (see Sec. 3.3.6), or by path *backtracking* (Kirbas and Quek, 2002, 2003; Quek and Kirbas, 2001; Stefancik and Sonka, 2001). The latter approach is closely related to minimal path methods from Sec. 3.3.4.

As greedy processes relying on low-level binary criteria, region-growing techniques heavily depend on the robustness of the underlying features. To cope with potential under- and over-segmentation issues, some works have implemented online schemes to dynamically adjust parameters such as inclusion thresholds (Larralde et al., 2003; Lorenz et al., 2003; Manniesing and Niessen, 2004; Masutani et al., 1998; Sekiguchi et al., 2005; Tschirren et al., 2005; Yi and Ra, 2003). Such adaptive schemes exploit the iterative, coherent nature of the propagation process. In (Lorenz et al., 2003; Sekiguchi et al., 2005) for instance, it is stressed that adapting parameters independently for each vascular branch is a key factor for accurate and robust extraction. Although typically of a heuristic nature, dynamically adaptive criteria are simple, powerful examples of active collaborations between the extraction process, the features and the underlying models.

3.3.3 Active Contours

Active contours constitute a popular class of image segmentation techniques that evolve an interface through different forces: *external* forces, derived from the image, and *internal*, model-based forces, constraining the contour geometry and its regularity (McInerney and Terzopoulos, 1996). In Sec. 3.2.3, we already mentioned the special case of active contours for 2D cross-section segmentation as an advanced example of a cross-section feature. For direct 3D extraction of vascular structures however, classical implementations of active contours, *e.g.* using curvature-based regularization, are usually unfit as they prevent the capture of thin, elongated surfaces. Hereafter, we present refinements specifically targeting vessel-like objects.

Parametric Active Contours Parametric active contours (Kass et al., 1987) rely on an explicit Lagrangian formulation of a contour evolution. Their advantages are their computational efficiency and ease of implementation in 2D. In 3D however, their param-

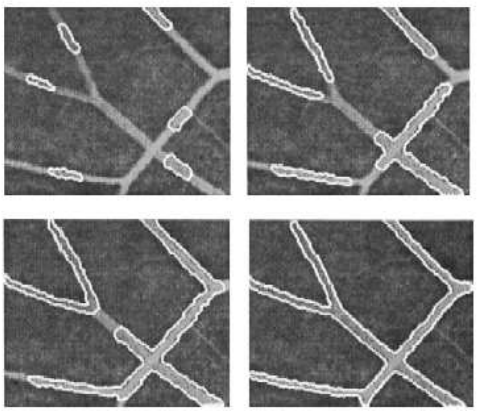


Figure 3.16: Topology adaptive snakes (*t-snakes*) from (McInerney and Terzopoulos, 1999). From top to bottom, left to right: evolution on retinal angiography using multiple merging *t-snakes*. Illustration based on material from (McInerney and Terzopoulos, 1999).

eterization becomes more complex. The reader may refer to (Delingette and Montagnat, 2001) for a comprehensive overview.

Topology-adaptive snakes (t-snakes) were introduced in (McInerney and Terzopoulos, 1999, 2000) (see Fig. 3.16). In contrast to classical snakes, contour splitting and merging can be controlled. The authors claim that this control over topology, combined with a specific re-parameterization scheme, allows for a better capture of thin and branching structures.

Eigen-snakes from (Toledo et al., 2000b,a) were specifically designed for vascular segmentation. Their energetic formulation exploits the direction of the target vessels, estimated through a principal component analysis on the distribution of gradient vectors.

Montagnat (1999) developed a 3D parametric active surface technique, where the evolution is constrained axially with respect to an existing centerline curve. Similarly, Yim et al. (2001) introduced a deformable tubular model, where a surface mesh is optimized with respect to the centerline curve, using a tubular coordinate system. Particular care is given to the parameterization schemes, avoiding for instance self-intersection artifacts. The work of (Frangi et al., 1999a) can also be considered as parametric active contours, as the B-spline tensor surface modeling the vessel surface is optimized thanks to explicit control points. Subsequent refinements can be found in (Frangi et al., 2000), with a more local, incremental optimization scheme. Finally, Mille et al. (2008) proposed a 2D, parametric deformable model evolving a centerline curve with varying radius. This technique relies notably on a novel local region-based energetic formulation, and was recently extended to 3D in (Mille and Cohen, 2009).

Implicit Active Contours Implicit active contours have become particularly popular for image segmentation, especially through the development of *level-set* techniques. They rely on an Eulerian formulation of contour evolution through partial derivative equations. The contour is embedded as the zero level of a higher dimension (level set) function. The level-set framework does not suffer from parameterization issues inherent to explicit techniques and topology changes are handled implicitly. For a more detailed discussion of these techniques, please refer to reference books such as (Angelini et al., 2005; Sethian,

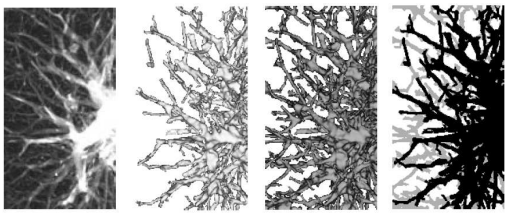


Figure 3.17: Co-dimension 2 level-set: *Curves* (Lorigo et al., 2001). From left to right: original image (CT lung scan of bronchi), result obtained by classical co-dimension 1 level-set, *Curves* segmentation and super-imposition of the results. Extra structures obtained by *Curves* are depicted in light gray. Illustration based on material from (Lorigo et al., 2001).

1996b, 1999).

Curves (Lorigo et al., 2001) is a vessel-dedicated *co-dimension 2* level-set scheme. Its core theoretical principle is to evolve a 1D curve on a 3D domain. In practice, the algorithm of (Lorigo et al., 2001) evolves a width-limited surface, called the ϵ*-level set*. A new energy term constrains the lowest curvature of this surface, assumed to be the vessel's principal direction. This approach has proved to be well adapted to vascular segmentation, as depicted in Fig. 3.17, and has been subsequently used in (Luboz et al., 2005).

Higher order active contours (Rochery et al., 2005a, 2006) introduce new geometric constraints on the contours, expressed as non-local quadratic forces modeling intricate contour interactions. In particular, the forces introduced in (Rochery et al., 2005a, 2006) penalize local widening of the contour and repulse proximate branches. Examples of evolution patterns are given in Fig. 3.18. The method was applied in 2D on satellite images for road extraction and on 2D angiograms. Its extension to 3D raises considerable discretization and computational issues. Another contribution of this work is the choice of *phase fields* over classical level-set implementations (Rochery et al., 2005b) for neutral initializations. The *ball measure* from (Nain et al., 2004) follows similar ideas. It was proposed as a soft shape prior for classical level-set evolution, penalizing local widening of the contour (see Fig. 3.19).

Some recent developments have focused on variational formulations of *flux maximization* for the extraction of elongated structures. Flux maximization techniques aim at aligning the surface normals to a data-based vector field (typically, the gradient vector field). Such an idea was first applied to vascular segmentation in (Vasilevskiy and Siddiqi, 2002) with the derivation of a multiscale gradient flow (see Fig. 3.20), and was later extended to multiscale Hessian-based flow in (Descoteaux et al., 2004). Following similar ideas, the *edge alignment* flow from (Kimmel and Bruckstein, 2003) was later combined in (Holtzman-Gazit et al., 2006) with Chan and Vese's minimal partition model (Chan and Vese, 2001), in a hierarchical segmentation framework. The use of the minimal partition model notably ensures global optimality of the energy minimization. The so-called *capillary* force from (Yan and Kassim, 2005) is another edge alignment term proposed for level-set evolutions. Authors report increased accuracy over *Curves* from (Lorigo et al., 2001). All these methods exploit vectorial information and share variational formulations closely related to anisotropic filtering techniques from Sec. 3.3.1.

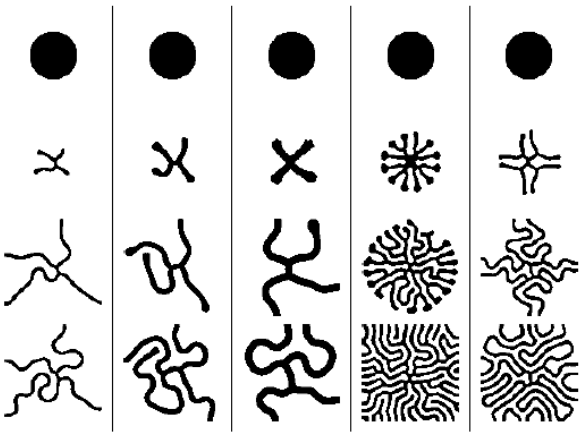
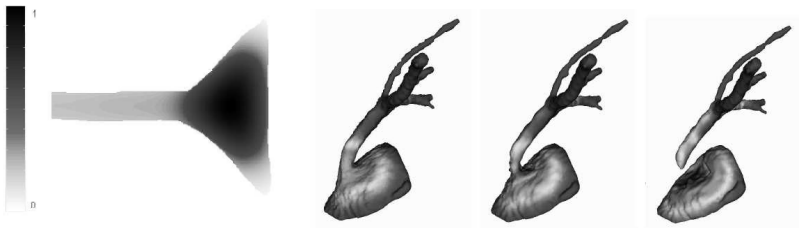


Figure 3.18: Higher order active contours from (Rochery et al., 2005a).
Examples of geometric evolutions without data convection terms, with quadratic forces constraining the contour to tubular-like patterns. Illustrations for different sets of geometry-controlling parameters, based on material from (Rochery et al., 2005a).



(*a*) (*b*)

Figure 3.19: Ball measure from (Nain et al., 2004).
(a) Ball measure on a 2D example: darker pixels indicate higher measurements. (b) Application to coronary CTA. From left to right: evolution of a level-set extraction using the ball measure. The ball measure penalizes widening and progressively disconnects the artery from the heart chamber into which the initial segmentation leaked. Illustrations based on material from (Nain et al., 2004).

Active contours enjoy a flexible, general framework that allows the integration of various internal and external forces. Practical issues may come from classical gradient-based Euler-Lagrange optimization schemes, which converge to local minima of the energetic formulation. The optimization landscape may contain numerous spurious minima, so that in practice, some active contour techniques require initializations close to the final result. The use of vessel-dedicated features and forces, coupled with robust optimization schemes such as (Kimmel and Bruckstein, 2003) help lessening such requirements. Initialization constraints can be further relaxed by exploiting the implicit handling of topology changes in level-set techniques. For instance, one can use isolated seeds along the vessels that will eventually reconnect (see Fig. 3.17). False positive seed regions are also tolerated to a certain point as they are expected to disappear under contour regularization. Implicit,

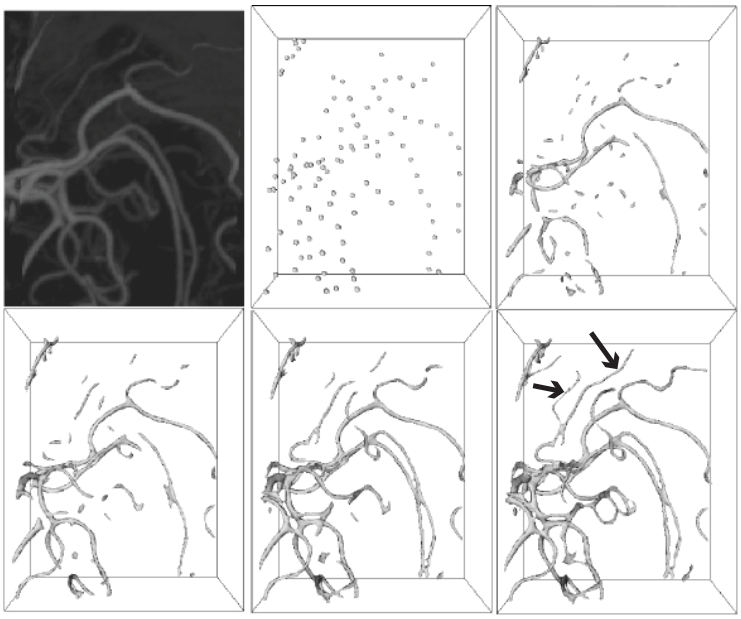


Figure 3.20: Flux-based level-set from (Vasilevskiy and Siddiqi, 2002). From top to bottom, left to right: original image (detail of 3D computed rotational angiography of the head), initialization, successive evolution steps until convergence. Notice the capture of thin vessels (arrows). Illustration based on material from (Vasilevskiy and Siddiqi, 2002).

thus uncontrolled topology changes may as well become a liability, causing segments to spuriously merge or disconnect. Following the same observations, branchings are usually tolerated, but not explicitly detected.

3.3.4 Centerline-Based Methods

The aforementioned region-growing and active contour methods (Sec. 3.3.2 and 3.3.3) aim at explicitly and directly detecting vessel contours, generally using low-level operations such as pixel-wise inclusion criteria. *Centerline-based* techniques focus on directly extracting the vessel centerline (see Fig. 3.21). Arguably, such methods are not true lumen segmentation techniques, as complete volume information is generally not part of the final result. As emphasized in (Aylward and Bullitt, 2002), such techniques seek additional robustness by relying on higher-level information, such as the localization of the center of the vessel, the estimation of its direction and scale. In most cases, a rough volume segmentation can thus be obtained directly by enriching the centerline with the underlying scale information. The extracted centerline can be used to efficiently constrain an accurate segmentation of the contours (see Sec. 3.3.6).

Direct Centerline Tracking Most centerline tracking algorithms are semi-automatic approaches, requiring as an input a seed point at the root of the vascular tree. Tracking progresses along the vessel by iterative *prediction* and *correction* steps, following schemes classically used for the statistical estimation of dynamic processes (Kitagawa and Gersch, 1996; West and Harrison, 1997).

The prediction of successive centerline positions is generally performed with respect

to estimations of the vessel direction, which may come from model-based and/or image-based sources: geometric models (Sec. 3.1.2) and/or image features (Sec. 3.2.2). The simplest model-based prediction scheme is to reuse directly the previous direction (Tek et al., 2001; Tyrrell et al., 2007). For increased robustness, filtering of the position history can be employed (Flasque et al., 2001; Wink et al., 2000a). *Kalman filtering*, which is a theoretically optimal prediction and correction scheme under Gaussianity and linearity assumptions [4], has also been applied to vascular tracking (Flórez Valencia et al., 2006; Gong et al., 2003; Wörz and Rohr, 2004, 2007, 2008). The local orientation of the vessel can be estimated from the data thanks to local geometry features (Sec. 3.2.2). Hessian-based estimation was used in (Aylward and Bullitt, 2002), moments of inertia in (Boldak et al., 2003; Hernàndez Hoyos et al., 2000; Hernàndez Hoyos, 2002; Hernàndez Hoyos et al., 2005, 2006; Larralde et al., 2003; Reuze et al., 1993), and gradient vector distribution in (Agam and Wu, 2005; Agam et al., 2005).

To ensure robust tracking, predicted positions are generally corrected, *i.e.*, recentered. Sphere-based geometric recentering schemes have been proposed in (Carrillo et al., 2005) and (Carrillo et al., 2007; Hernàndez Hoyos et al., 2006), relying on local thresholding and inertia moments, respectively. Recentering can also be performed by a local optimization of medialness features. Simple examples are found in (Verdonck et al., 1995; Wesarg and Firle, 2004), where the predicted centerline position is recentered as the center of mass of cross-sectional contour points detected along 1D rays. 2D cross-sectional features (Sec. 3.2.3) are particularly popular in this context (Aylward and Bullitt, 2002; Fridman et al., 2003; Fridman, 2004; Niessen et al., 1999; Tek et al., 2001; Verdonck et al., 1995; Wink et al., 1998, 2000a). Applied on a local neighborhood (Aylward and Bullitt, 2002; Fridman et al., 2003; Wink et al., 2000a), recentering schemes have proved to be robust to predictions lying outside the vessel lumen (Wink et al., 2000a). Correction can be further refined with 2D active contours performing an accurate segmentation of the cross-section (Lee et al., 2007) (see Sec. 3.3.6). As an alternative to 2D cross-sectional recentering, correction by local optimization of 3D models was used in (Friman et al., 2008a; Tyrrell et al., 2007; Wörz and Rohr, 2008).

Direct tracking schemes explore the image very sparsely, reaching levels of time and memory performance suitable for large 3D datasets. The explicit embedding of geometric models generally results in improved robustness over previously mentioned methods. Nonetheless, they remain local optimization processes prone to premature stopping in presence of anomalies such as stenoses and aneurysms. In practice, some of these methods require a fair amount of user interactivity (Aylward et al., 1996; Larralde et al., 2003). A solution to further improve robustness is the use of *multi-hypotheses* frameworks such as stochastic particle filters (see Sec. 3.3.5). A deterministic alternative was recently proposed in (Friman et al., 2008a), with a multi-hypothesis algorithm tracking in parallel the successive local maxima of a 3D template fitting process. Substantial gain in robustness is obtained with the increase of the depth of the search tree.

Most centerline tracking techniques follow one branch at a time, relying on manual reseeding to extract a complete tree. To automatize bifurcation handling, some authors have proposed to perform the segmentation of the lumen locally and rely on topological criteria, similarly to aforementioned region-growing and wave propagation schemes (Sec. 3.3.2). For instance, the techniques from (Carrillo et al., 2005, 2007; Flasque et al., 2001) perform the analysis of connected components on local spheres and parallelepipeds. Due

4. Such assumptions may not hold in practice because of sudden changes in curvature along the vessel. Particle filters are alternative schemes dealing with non-linear, non-Gaussian processes (see Sec. 3.3.5).

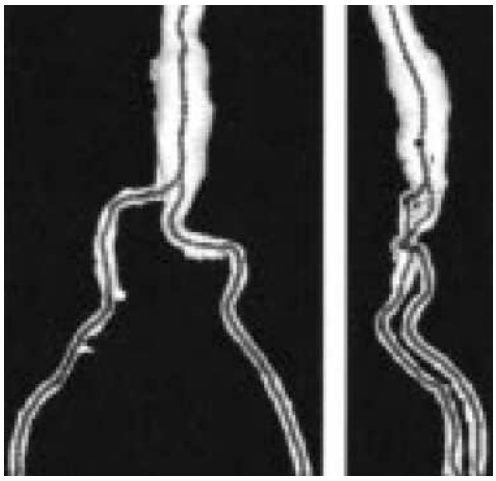


Figure 3.21: Direct centerline tracking (Wink et al., 2000a). Results samples on the ascending aorta and iliac branches from CT angiography. Illustration based on material from (Wink et al., 2000a).

to the use of local binary criteria, these approaches may suffer from the same robustness issues as region-growing techniques. In (Aylward and Bullitt, 2002; Fridman et al., 2003), online branching detection is performed through specific ridge-based features, without the need for a binary segmentation. An *a posteriori* branch detector, reportedly preliminary and not entirely satisfactory, is proposed in (Hernàndez Hoyos, 2002; Hernàndez Hoyos et al., 2005). In (Wink et al., 2000a), the tracking process is seeded from the smallest branches of the vascular tree and a specific pattern of a medialness filter is exploited to handle branchings. If a robust termination criterion is available, complete tree tracing can theoretically be achieved with automatic reseeding (Tyrrell et al., 2007).

Model-Based Optimization Between Two Points If start and end points are available for a given vascular segment, the centerline extraction problem can be advantageously constrained. A first approach to solve such a problem consists in optimizing a centerline model (see Sec. 3.1.2) with fixed boundary conditions. B-spline (Frangi et al., 1999a, 2000) (see Fig. 3.2) and cardinal spline (Wong and Chung, 2006) deformable models have been used in that context. Such methods balance data and model fitting through the parameterization of their energetic formulation. Optimization is classically performed by gradient descent schemes, which may be prone to local minima issues. In practice, such effects are mitigated by the additional robustness brought by the model.

Minimal Path Techniques Certain energetic models for two-point centerline extraction can be formulated as the minimization of a cumulative, monotonic cost metric integrated along the centerline path. The continuous formulation of minimal path optimization can be shown to correspond to a special case of geodesic active contours (Cohen and Kimmel, 1997). Efficient dynamic programming schemes are known to solve such problems *globally*. Minimal path approaches are particularly popular for centerline extraction. The cost metric is designed to favor centerline locations, which can be done for instance

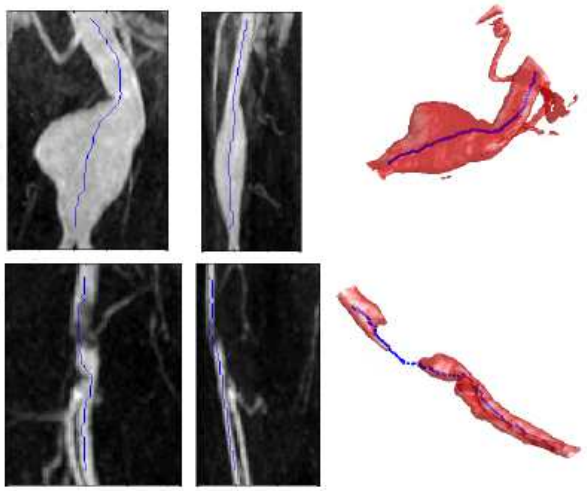


Figure 3.22: Shortest path approach: robustness to local anomalies. Top row: 2D and 3D views of a vessel with an aneurysm and overlayed backtracked centerline. Bottom row: 2D and 3D views of a severely stenotic vessel and overlayed backtracked centerline. Illustration based on material from (Lin, 2003).

through the reciprocal of a vessel-dedicated feature [5]. Global optimality is a key motivation for such techniques, resulting in good robustness even in cases of corrupted data or severe anomalies (see Fig. 3.22).

Minimal path techniques differ notably in the numerical optimization schemes they employ. Discrete, L_1 path optimization can be carried out with Dijkstra-like (Dijkstra, 1959) graph-based schemes (Gulsun and Tek, 2008b; Olabarriaga et al., 2003; Wink et al., 2004, 2001, 2000b). Such algorithms have generally the advantage of high computational efficiency, but can be shown to introduce metrication errors due to their discrete nature (Deschamps, 2001). The fast-marching algorithm (Adalsteinsson and Sethian, 1995; Sethian, 1996a; Tsitsiklis, 1995) is an optimization scheme which can be shown to be consistent with the continuous formulation of the minimal path problem (Deschamps, 2001; Cohen and Kimmel, 1997). Fast-marching optimization approximates the Euclidian (L_2) cumulative cost, yielding sub-voxel accurate paths. A counterpart, compared to discrete L_1 techniques, is a slight increase in computational cost due to more complex update schemes. Among other works, fast-marching optimization has been applied to vascular centerline extraction in (Avants and Williams, 2000; Deschamps and Cohen, 2001; Law and Heng, 2000; Lin, 2003; Young et al., 2001).

Closely related to the aforementioned minimal path approaches, *fuzzy connectedness* techniques (Udupa et al., 2002) (see also Sec. 3.3.6) correspond to the optimization of the L_∞ norm of a connectivity cost along discrete paths. In other words, the total cost of a path is that of its worst edge (lowest degree of connectivity). The *ordered region-growing* algorithm of (Choyke et al., 2001; Yim et al., 2000, 2003) follows the same principles. The ordered growing pattern defines an acyclic graph composed of the minimal L_∞ paths from the seed point to all the other points (see Fig. 3.23). Vessels can then be extracted by

5. A popular feature in this context is the Hessian-based *vesselness* measure from (Frangi et al., 1998), used for instance in (Olabarriaga et al., 2003; Wink et al., 2001, 2004).

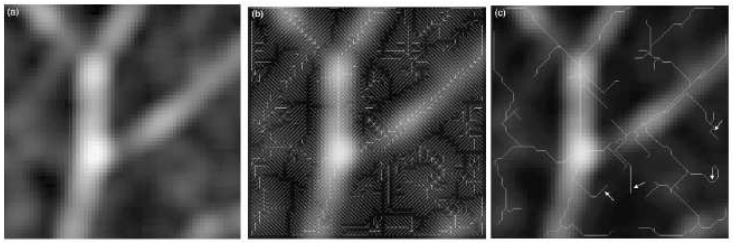


Figure 3.23: Ordered region-growing (ORG) method from (Choyke et al., 2001; Yim et al., 2000, 2003). (a) original image; (b) initial acyclic ORG graph superimposed; (c) graph obtained after trimming process. Several spurious branches are indicated with arrows. Illustration based on material from (Yim et al., 2003).

interactively specifying end points[6]. L_∞ path optimization can be performed efficiently, but it values paths in a non-cumulative fashion, which can result in the extraction of fairly irregular paths in practice.

Although algorithmically efficient, minimal path methods can induce the exploration of a large portion of the search space in order to extract the optimal centerline path. Bidirectional propagation, *i.e.*, computing paths from the start *and* end points simultaneously, considerably lowers computational costs (Deschamps and Cohen, 2001; Olabarriaga et al., 2003; Wink et al., 2000b). It is used in (Wink et al., 2000b) in conjunction with heuristic A* search to further reduce the search space. Another possible refinement is the *freezing* scheme from (Deschamps and Cohen, 2002a), which consists in prevents paths with too high cumulative costs from being propagated further[7]. Lin (2003) combines the freezing scheme with *anisotropic* versions of the fast-marching algorithm to greatly reduce the amount of exploration (see Fig. 3.24).

Classical issues with minimal path techniques include the extraction of eccentric or even erroneous *shortcut* paths (see (Li and Yezzi, 2006, 2007) and Fig. 3.25). These practical problems come principally from the cumulative nature of the optimization. They are alleviated in practice by the use of adequate, discriminative cost metrics. Strategies to reduce centering problems *a posteriori* can be found in (Deschamps and Cohen, 2001, 2002a).

Classical vessel-dedicated minimal path algorithms focus solely on extracting the centerline curve. Notable exceptions are the works from (Wink et al., 2004; Li and Yezzi, 2006, 2007; Benmansour et al., 2009; Benmansour and Cohen, 2009), which share the core idea of incorporating an additional dimension corresponding to the vessel radius. The spatial position of the centerline and its associated radius are then optimized conjointly. This idea was first applied to 2D vascular segmentation in (Wink et al., 2004), using the multiscale feature from (Frangi et al., 1998). It was extended to 3D in the *4D minimal path* technique from (Li and Yezzi, 2006, 2007). The authors proposed two potentials expressed on multiscale spheres, based respectively on similarity and contrast criteria. The methods from (Wink et al., 2004) and (Li and Yezzi, 2006, 2007) are implemented as relatively straightforward modifications of the original Dijkstra and fast-marching algorithms,

6. In (Choyke et al., 2001; Yim et al., 2000, 2003), authors focus on automatically cleaning the graph using trimming techniques (see also Sec. 3.3.6).

7. The *freezing* scheme is used in (Deschamps and Cohen, 2002a) to extract the vessel surface. This corresponds to its use as a binary stopping criterion in a *wave propagation* scheme (see Sec. 3.3.2).

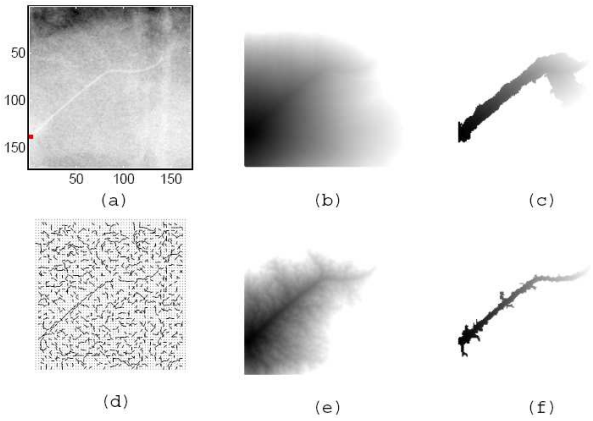


Figure 3.24: Refinements for fast-marching minimal paths (Lin, 2003). (a) Original image (start point in red). (b) Classical isotropic propagation. (c) Isotropic fast-marching with *freezing* from (Deschamps and Cohen, 2002a). (d) Orientation map yielding directional information. (e) Anisotropic fast-marching without freezing. (f) Anisotropic fast-marching with freezing. Illustration based on material from (Lin, 2003).

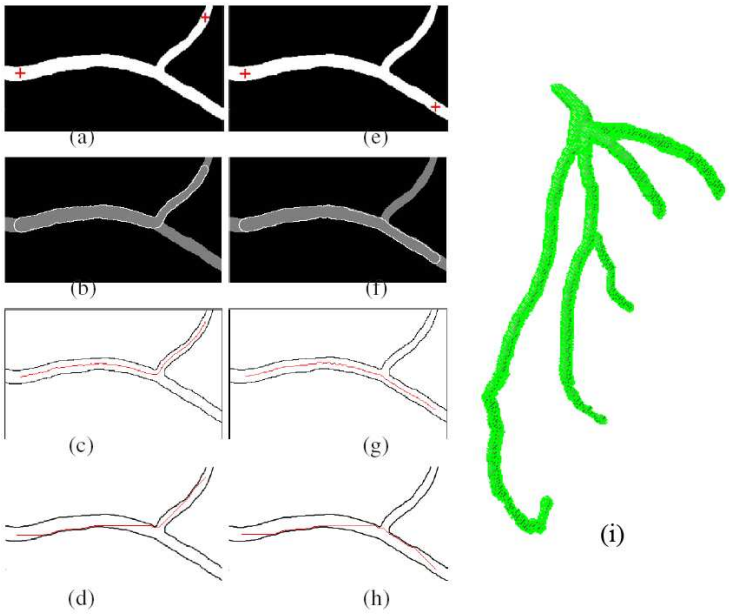


Figure 3.25: 4D minimal path technique from (Li and Yezzi, 2006, 2007). With local radius as an extra dimension, minimal paths are computed in 3D (resp. 4D) for 2D (resp. 3D) images. (a) and (e) 2D test images with start and end points; (b) and (f) surface approximation by the proposed method (envelope of the local circles/spheres); (c) and (g) centerline extracted by the proposed method; (d) and (h) illustration of the shortcut effect occuring with classical minimal path techniques; (i) 4D paths extracted from a 3D cardiac CTA, forming the coronary tree result. Illustrations based on material from (Li and Yezzi, 2006).

respectively. As important benefits, resulting paths are better centered and the vessel surface can be directly approximated by the envelope of the optimized circles/spheres (see Fig. 3.25).

Besides radius optimization, some works focused on exploiting directional information in minimal path frameworks. We already mentioned the development of anisotropic fast-marching flows in Lin (2003) (Fig. 3.24) favoring propagation in certain orientations. This idea was recently coupled with multiscale optimization in (Benmansour et al., 2009; Benmansour and Cohen, 2009), with a particular focus on the design of an adequate multiscale anisotropic metric and on the numerical optimization scheme. Authors in (Benmansour et al., 2009; Benmansour and Cohen, 2009) rely on the multiscale optimally oriented flux from (Law and Chung, 2008) as an underlying feature. Another approach to directional minimal paths is to consider the orientation domain as an additional dimension to be optimized, similarly to the aforementioned multiscale techniques. To the authors' knowledge, the application of this idea to angiographic data was first discussed in (Deschamps, 2001). The recent work of Pechaud et al. (2009) combines multiscale and orientation optimization by propagating 4D (space+scale+orientation) paths on 2D images. In particular, the authors demonstrate the capacity of their method to disambiguate overlap issues in 2D angiograms. The principal drawback of such advanced techniques is their relatively high computational cost, potentially prohibitive for 3D applications.

Minimal path techniques are commonly employed in interactive frameworks, requiring the definition of start and end points for each target vessel. Some works have proposed the definition of termination criteria to automatically stop the path propagation and relax the need for end points. Such criteria can for instance be devised through heuristic thresholds on the underlying features (Gulsun and Tek, 2008b). An alternative proposal is the analysis of *pioneer* points, defined by the first paths crossing a certain geodesic distance (Wink et al., 2001). Finally, Deschamps and Cohen (2002a) proposed to exploit the increase rate of the geodesic distance to stop the propagation. Given a stopping criterion, complex vascular trees can be extracted *a posteriori*, at convergence, by backtracking multiple minimal paths to a single proximal seed point (Gulsun and Tek, 2008b). A related approach can be found in the geodesic voting technique of (Rouchdy and Cohen, 2008), recently applied to 3D vascular trees in (Mille and Cohen, 2009). This method uses a measure of the *density* of minimal paths running through a given location to identify the main segments of the vascular tree. It relies on the observation that minimal paths from different end points converge and concentrate in highly anisotropic areas corresponding to true vessels. This technique can actually be applied to extract vascular trees from fully optimized action maps, without the need for a stopping criterion. Geodesic voting then serves as an *a posteriori* pruning criterion, keeping only locations of high path densities(Mille and Cohen, 2009).

3.3.5 Stochastic Frameworks: Particle Filters and Markov Marked Point Processes

In this section, we highlight two stochastic frameworks which were recently successfully adapted to vascular segmentation: *particle filtering* and *Markov marked point processes*. In both cases, seminal works showed especially promising results in terms of extraction robustness. We emphasize the overall elegance and versatility of these schemes. They benefit from well-established Bayesian theoretical frameworks and enable high-level designs, seamlessly integrating model- and data-based information.

Particle Filters *Sequential Monte-Carlo* techniques, also known as *particle filtering*, are recursive Bayesian estimation schemes which can be used, among other applications, to track the *posterior* probability distribution of a dynamic process (Arulampalam et al., 2002; Doucet et al., 2001). We already mentioned Kalman filtering as the optimal tracking scheme for linear Gaussian state-space problems. Unfortunately, such simple assumptions generally do not hold for vessel characteristics (sudden scale variations, direction changes, presence of bifurcations). Instead of relying on a deterministic estimation scheme, particle filters handle *non-linear* processes through *stochastic*, population-based sampling schemes. Probability densities are approximated by non-parametric *Monte-Carlo* methods, relying on a discrete population of weighted samples, the so-called *particles*.

Particle filtering can be adapted to vascular segmentation by considering the tracking task as a recursive estimation problem. Particle filter-based techniques can thus be considered as direct tracking approaches (Sec. 3.3.4). The key difference with methods from Sec. 3.3.4 is that particle filters employ a stochastic, multi-hypothesis approach. The particles form an approximation of the *entire* probability distribution on the state-space, which consists of the model parameters (*e.g.*, vessel location, scale, orientation, appearance parameters) to be optimized. We emphasize the versatility of the state-space definition, enabling the use of high-level, specialized models and features. The particle population is evolved by predicting new model parameters, obtained by sampling proposal distributions. Such distributions typically incorporate prior knowledge on the process dynamics (scale and direction changes, for instance). Subsequent correction is notably performed by evaluating the parameters likelihood on the data.

Particle filters were first applied to vessel segmentation in (Florin et al., 2005) (see Fig. 3.26), using an elliptic cross-section model. Their likelihood model consists of a contrast-based term, evaluated by a so-called *ribbon* measure (see Sec. 3.2.3) and a region-based statistical mixture on intensity distributions. This seminal work proved that particle filters are less sensitive to premature failures than mono-hypothesis tracking algorithms, being more robust to noise and local anomalies. Refinements were later proposed in (Florin et al., 2006), with the use of circular shortest paths (Appleton and Sun, 2003) to determine cross-section areas.

The main counterpart to the robustness of particle filters is their computational cost, proportional to the number of particles. With potentially very large state-spaces, a relatively high number of samples has to be maintained. Also relying on Monte-Carlo probabilistic tracking, the work from (Schaap et al., 2007a) introduces a new appearance model for vessels with heterogeneous surroundings, combined with a tubular geometric model. A subsequent work from the same authors (Schaap et al., 2007b) details algorithmic optimizations coping with redundant computations. The work from (Lesage et al., 2008) proposes a new sampling scheme, based on a constrained medial-based geometric model. It discusses the efficient construction of a discrete representation of the posterior distribution, from which a relatively low number of samples is required to ensure tracking robustness.

Particle filters are theoretically able to handle bifurcations implicitly as multi-modal distributions (splits of the particle population, see Fig. 3.26). In practice, sample *impoverishment* (Arulampalam et al., 2002) usually causes secondary branches to be lost (Florin et al., 2005, 2006). In (Florin et al., 2005, 2006), explicit detection is performed using K-means clustering. Each detected cluster is re-populated and evolved independently. The recent work from (Allen et al., 2008) also employs K-means to separate the propagation in each branch, but additionally proposes the use of Markov Chain Monte-Carlo rejuvenation

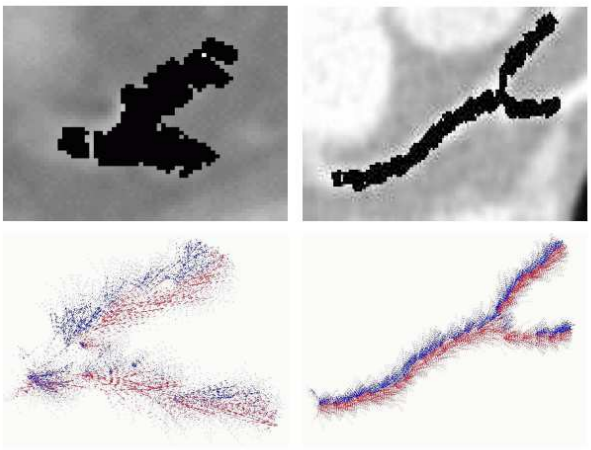


Figure 3.26: Particle filtering from (Florin et al., 2005). Application to the tracking of coronary arteries from cardiac CTA. Top row: particle mean states are overlaid on the original image. Bottom row: particle distributions split at bifurcations. Such a situation is handled with clustering and repopulation methods. Illustration based on material from (Florin et al., 2005).

to increase robustness to branchings.

Finally, it is worth highlighting the existence of deterministic algorithms also based on the idea of multi-hypothesis tracking. For instance, the method of (Friman et al., 2008a) is a fast, robust tracking scheme based on 3D template matching (see Sec. 3.1.3 and 3.2.2). The authors demonstrated considerable gains in robustness obtained from their multi-hypothesis approach.

Markov Marked Point Processes Throughout this document, we emphasized the importance of using high level information on the vasculature geometry and appearance. Unfortunately, advanced models often come with numerous parameters, inducing very large search spaces and complex optimization landscapes. Some of techniques previously detailed, such as particle filtering, are able to cope with high-dimensional spaces. On the other hand, they still rely on local optimization schemes and, generally, on user interaction for their initialization. A promising alternative is found in the fully-automatic, globally optimal framework offered by *Markov marked point processes*.

Markov marked point processes were used in (Lacoste et al., 2005) to extract road networks from remote sensing images. This seminal work was later adapted to the segmentation of coronary arteries from 2D angiograms in (Lacoste et al., 2006) (see Fig. 3.27). Vessels are modeled locally as piece-wise linear segments of varying locations, lengths, widths and orientations. Lacoste et al. (2006) use simple contrast measures defined on these segments. The extraction process proceeds as follows. A population of segments is initialized randomly on the image and is evolved through a *reversible jump Markov chain Monte-Carlo* scheme. This algorithm governs the evolution of the parameters, along with the *birth* and *death* of segments. A *simulated annealing* scheme ensures the convergence to a clean and stable structure that corresponds to the vascular tree. The initial result is refined through a similar, but more accurate hierarchical *polyline* process focusing on

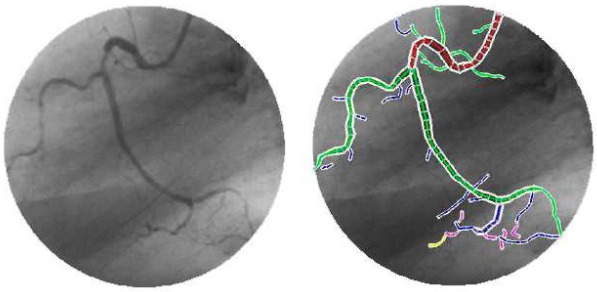


Figure 3.27: Markov marked point processes (Lacoste et al., 2006). Application to 2D coronary angiography. Left: original image. Right: extracted segments. Illustration based on material from (Lacoste et al., 2006).

border information to reconnect and extend the vascular network.

This approach is particularly appealing in the way prior knowledge on geometric primitives and on their *interactions* (connections and angles between segments, repulsive forces between proximate parallel branches) is elegantly expressed within an homogeneous, well-posed theoretical framework. This additional high-level knowledge, coupled with the flexibility of the stochastic process, is the key ingredient for both robust and accurate results. So far, this approach has only been applied on 2D data. Although there is no theoretical obstacle to its extension to 3D, we expect its already very high computational cost to increase significantly. This makes such a technique unfortunately prohibitive by today's clinical standards.

3.3.6 Post-Processing

Initial extraction results may be lacking in different aspects: surface information may be missing or inaccurate, the vessel's topology may be incorrect, non-vessel regions may be included, vessel segments may be disconnected or missing. Hereafter, we briefly mention post-processing steps that were proposed to cope with such issues.

From Centerlines to Contours Most centerline-based techniques (Sec. 3.3.4) extract the centerline without providing an accurate surface segmentation. The availability of the centerline structure greatly eases subsequent surface extraction through geometric and spatial constraints.

A first approach is to segment the vessel wall iteratively in cross-sectional planes. Among cross-section-dedicated techniques (Sec. 3.2.3), ray-casting schemes, 2D active contours (see also Fig. 3.28) and circular dynamic programming schemes are good candidates for the accurate segmentation of cross-sectional contours.

From a collection of 2D contours, a 3D surface can be optimized using geometric models such as *generalized cylinders* (O'Donnell et al., 1994) and *finite element* techniques(O'Donnell et al., 1998) (see also Sec. 3.1.2). Another possibility is to extract or refine the surface directly in 3D: 3D parametric active surfaces are used in (Montagnat, 1999; Yim et al., 2001; Wong and Chung, 2006; Mille and Cohen, 2009), level-set evolutions in (van Bemmel et al., 2003), B-spline tensor surfaces in (Frangi et al., 1999b,a, 2000) (see Fig. 3.2). Another alternative, recently applied successfully to CTA coronary

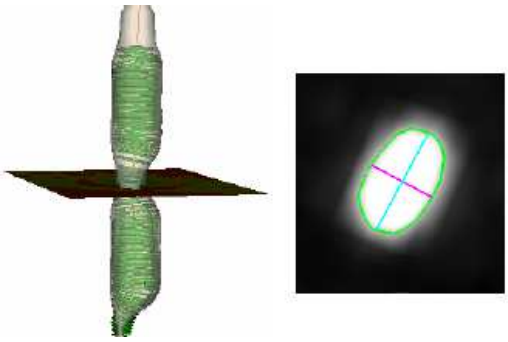


Figure 3.28: Parametric active contours for 2D cross-sections (Hernàndez Hoyos, 2002; Hernàndez Hoyos et al., 2005). Given the vessel centerline, lumen segmentation is performed thanks to 2D snakes optimized in perpendicular cross-sections. Illustration based on material from (Hernàndez Hoyos, 2002; Hernàndez Hoyos et al., 2005).

arteries, can be found in (Schaap et al., 2009b). It combines graph-cut optimization and kernel regression to obtain an accurate segmentation of the coronary lumen. A particularly interesting refinement introduced in this work is the exploitation of the intensities along the pre-extracted centerline to derive a local similarity criterion to valuate graph-cut edges.

From Contours to Centerlines Recovering centerline information from a 3D binary segmentation is generally formulated as a *skeletonization* problem (Buhler et al., 2002; Suri et al., 2002b). Most skeletonization algorithms follow a *homotopic thinning* design (Malandain et al., 1993; Palagyi and Kuba, 1998), in which points are removed from the binary object in a particular order. Distance transforms are classically used (Metz et al., 2007; Pudney, 1998; Puig et al., 1997). Gradient flux maximization is a more recent alternative (Bouix et al., 2005; Siddiqi et al., 2002). Pruning strategies can be employed to remove spurious branches (Boskamp et al., 2004; Bouix et al., 2005; Wan et al., 2000, 2002b) and centerline models can be applied for subsequent smoothing of the thinning result (see Sec. 3.1.2).

Spatial Coherence and Topology Low-level pixel-wise techniques are prone to issues such as holes, isolated false positives and noisy contours. Markov random fields regularization (Chung et al., 2004; Wong et al., 2004), maximum intensity projection analysis (Gan et al., 2004) or anisotropic smoothing (Yang et al., 2004) (see Fig. 3.13), help correct such problems. An alternative is the use of mathematical morphology: hole-filling is performed in (Wan et al., 2000, 2002b), *regulated morphology* operators are used in (Agam et al., 2005; Wu et al., 2004).

Region Competition and Splitting If an initial segmentation result includes non-vessel parts, region competition can be used to disconnect them from the actual vessel. We already mentioned *competitive region-growing* (Yi and Ra, 2003) and the hierarchical segmentation scheme from (Holtzman-Gazit et al., 2006) in Sec. 3.3.2 and 3.3.3, respectively. Another candidate technique is *fuzzy connectedness* and its various refinements (Udupa et al., 2002), successfully applied to artery-vein separation in (Lei et al., 2001).

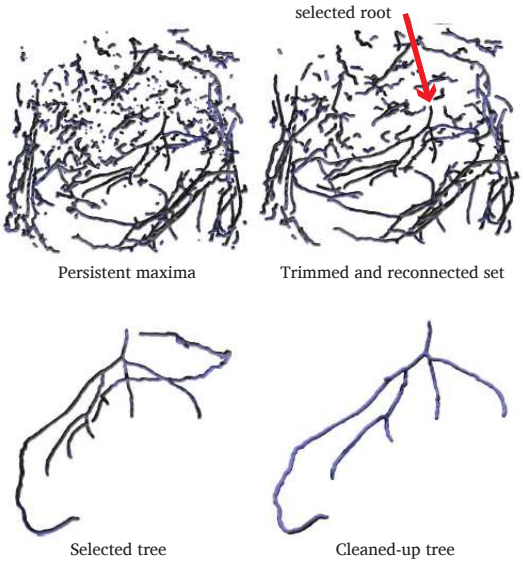


Figure 3.29: Graph-based workflow for coronary segmentation from cardiac CTA (Szymczak et al., 2005). Initial point candidates are obtained as local robust maxima of intensity. Heuristics are used to trim the graph and reconnect disconnected branches. The target vascular tree is extracted as a minimum spanning tree, given its user-provided root location. Subsequent cleanup is applied to obtain the final result. Illustration based on material from (Szymczak et al., 2005).

Graph-Based Clean-Up and Reconnection Graph techniques can also be used to clean-up cluttered segmentations, exploiting the intrinsic tree structure of vascular networks. A popular example is the use of minimum spanning tree algorithms, for which different valuation strategies have been proposed: Euclidian (Bullitt et al., 1999) or Mahalanobis distances (Jomier et al., 2005), energies accounting for intensity variations and path regularity (Florin et al., 2004). Various heuristics have been proposed to trim or reconnect centerline graphs (Lee et al., 2007; Szymczak et al., 2005; Yim et al., 2000, 2003) (see Fig. 3.29). In (Agam et al., 2005; Wu et al., 2004), such heuristics are given for a graph of so-called *fuzzy spheres*, with the additional notion of radius similarities to value graph edges. Closely related, minimal path techniques can also be used to reconnect disconnected regions (Cohen and Deschamps, 2001; Deschamps and Cohen, 2002b).

3.3.7 General Considerations on Extraction Schemes

The choice of an adequate extraction scheme depends notably on the degree of interactivity and type of output required, on the robustness, accuracy and computational efficiency needed. For instance, region-growing schemes are usually straightforward to implement, fast and minimally interactive. Their main issue is the design of robust inclusion and stopping criteria, which may be difficult in practice. Two-point centerline-based minimal path techniques have earned their popularity from their robustness, computational efficiency and control over boundary conditions. Counterparts are the increased need for manual interaction and, in classical implementations, the lack of surface segmentation. As emphasized throughout the document, good practice can be to *combine* extraction

schemes. For instance, 1D centerlines can first be obtained from minimal paths and serve to initialize active contours for an accurate segmentation of the vessels' surface (see Fig. 3.11).

3.4 Discussion

The medical interest in 3D vascular segmentation and the challenges it raises have motivated a tremendous amount of research work. Techniques from almost all image processing frameworks have been explored. The complexity of the methods proposed in the literature depends heavily on the difficulty of the application and on the levels of automation, accuracy and computational efficiency targeted. Angiographies with prominent vessels and homogeneous backgrounds might be successfully handled with simple schemes we classified as pre-segmentation techniques (Section 3.3.1) while most complex ones can involve numerous entangled refinement steps. In practice, proposed methods differ so greatly in their applicative frameworks and purposes that direct comparisons are often meaningless. Instead, the axes developed in this review aim at clarifying the key design issues that led to the proposed techniques:

- What vascular-specific prior knowledge is expressed?
- How is it exploited, measured? How reliable are these measurements?
- How are these measurements controlled, optimized, corrected, or used to adapt underlying models?

As was illustrated in Table 3.1 in the early part of this bibliographic study, existing methods are built upon a large variety of models, features and extraction schemes. By decomposing a whole framework into elementary components, one gains insight on its core properties, strengths and limitations. A first remark is that model-based information is often embedded *implicitly* in the features and extraction schemes. For instance, the L_2 minimization performed by fast-marching tracking (Section 3.3.4) expresses assumptions on the geometric smoothness of the centerline paths. Such concealed aspects must be taken into account to fully understand the practical behavior of a given segmentation framework.

A second key observation is that most advanced solutions rely on combinations of several models and corresponding features, even mixing several extraction schemes. They exploit vessel-specific prior knowledge such as the scale and elongation of the vessels of interest. Towards the design of new algorithms, one can explore new *sequential* combinations to obtain more efficient and more automated global schemes: pre-processing can be added to improve the quality of the image data, pre-segmentation to automatize the initialization of dedicated extraction schemes, rough delineation to constrain further refinement. More and more works in the literature consist in elaborate chains of existing components. At each step of the process, there usually exists a variety of possible choices. Choosing the right component then depends on computational constraints and, more importantly, on the prerequisites of the subsequent blocks. For instance, high requirements on the inputs of an extraction scheme limit the feasibility of automatic initialization.

Combinations can also be built from a *hierarchical* point of view, *i.e.*, by having components cooperating at different levels of abstraction. For instance, the framework of Lorenz et al. (2003) is organized in voxel, segment and tree levels. Such layered designs offer the valuable possibility of exploiting prior knowledge at the adequate level of granulometry. Local tubularity hypotheses might be expressed at the segment level, while the tree level controls higher-order assumptions on bifurcations and centerline regularity. An even more

striking example can be found in the recent work from Socher et al. (2008), where 2D coronary angiographies are segmented using a hierarchical *marginal space* learning paradigm trained by *probabilistic boosting trees*. Marginal spaces correspond to different levels of abstraction: border points are detected and grouped to form cross-segments, which are finally linked cross-segments into vessel pieces. The inter-dependency of the different components is even more intricate with such layouts. It is crucial to identify these dependencies as they induce potential performance bottlenecks. Improvements can be obtained by exchanging components for more robust ones, or by changing how critical pieces of information are obtained. For instance, cross-sectional features (Section 3.2.3) depend on the availability of the local direction information; the direction can be estimated by other local features (risks of unreliability), by evaluation over a set of possible directions (potentially costly) or can be controlled at a higher level, *e.g.* predicted by a tracking extraction scheme.

Besides combinations of techniques, an important trend is the development of *adaptive* extraction schemes, underlining complex dynamics between models and features. Instead of relying on a fixed formulations to extract the entire vascular network, each vessel is considered as an individual target for which models, features and extraction parameters are tuned. This principle is particularly applicable to tracking methods, in a broad sense, from region-growing to particle filtering. By performing the segmentation as a dynamic process, such algorithms can exploit information acquired *during* the segmentation itself to adapt models and features, lowering under- and over-segmentation issues. Adaptivity is often expressed in a heuristic fashion, via *ad hoc* adjustments of local thresholds for instance. Some frameworks, especially Bayesian approaches (see particle filters and Markov marked point processes, Section 3.3.5), offer theoretically well-defined adaptive schemes.

Aiming for a generic, flawless technique is probably illusory. As stressed previously, the adequacy of a given technique is highly dependent upon one's practical requirements. The same solution might be overkill for simple applications while being insufficient for others. Additionally, specifically tailoring an algorithm towards one's applicative needs is generally unavoidable in order to reach high levels of efficiency. We believe, however, in the general principle of building more efficient and sophisticated collaborative systems out of simple components. To that aim, we again emphasize the importance (and difficulty) of robustly incorporating prior knowledge at the adequate abstraction level. In that regard, probabilistic frameworks such as particle filters and Markov marked point processes are particularly attractive in their capacity to handle complex object parameterizations (hence building abstractions) while maintaining well-defined schemes to merge model-based prior information and data likelihood derived from image features. As algorithmic designs tend to become always more intricate, a deep understanding of the properties of each core component remains essential.

Chapter 4

Geometric and Appearance Models

In this chapter, we present our geometric model for vascular structures. It provides us with a compact representation making explicit the key geometric properties of the vessels of interest, namely their local size (radius) and orientation. It allows us to easily express constraints on the geometry of the recovered structures.

In addition to the geometric model, we analyze the *appearance* of the vessels in our modality of choice, cardiac CTA. In particular, we detail the assumptions that can be exploited in terms of typical intensity ranges. We also discuss the *radial* appearance of CT coronaries towards the choice of a suitable parametric model for the experiments of Chapter 5.

4.1 Discrete Medial-based Geometric Model

In this section, we first discuss some geometric modeling works related to our approach. For a more general presentation of existing models for vascular structures, the reader may refer to Sec. 3.1. Second, we present our discrete, medial-based geometric model. Our model encodes information about the centerline position, tangential direction and radius. It relies notably on a compact parameterization scheme coupling spatial positions and local directions.

4.1.1 Related Works

Our geometric model is largely inspired by medial-based representations and other vessel-dedicated centerline-based models. This section gives a brief overview of these seminal works.

Medial representations of general shapes In 2D, a *medial-based* representation aims at compactly encoding the core topological structure of a shape as a structured collection of medial axes (the "skeleton" of the shape). The *medial axis* of a 2D shape was mathematically defined in (Blum, 1967, 1973) as the centers of maximal inscribed disks (Fig. 4.1). A key property is that given the medial axes and the associated distance/radius mapping, the original shape can be reconstructed *exactly* in the continuous case (Blum, 1967, 1973). One can thus represent any 2D shape as a collection of disks. Blum named

this representation the *Medial Axis Transform* (MAT). A closely related medial-based representation is the Smoothed Local Symmetry (SLS) model from (Brady and Asada, 1984) which yields, in particular, a different behavior at branchings and axis end locations.

The MAT gives a mathematical definition of the medial structure of a shape. In other words, given a shape, then the location and number of axes and branching and end points are uniquely and strictly defined. The direct transposition of this mathematical definition to natural digitized shapes raises theoretical and practical issues. In particular, a recurrent criticism of the MAT is its instability to shape perturbations. Small protrusions and surface irregularities can significantly affect the MAT by altering the location of the axis and causing the apparition of spurious branches (see Fig. 4.1). In practice, methods relying on the original MAT definition (see Sec. 3.3.6) generally require subsequent correction and pruning schemes. Scale-space approaches can alleviate such issues (Naf et al., 1996; Pizer et al., 1998). More generally, there exist numerous variants of medial-based representations in the literature. They differ in their mathematical definitions and practical properties. The reader may refer to (Fridman, 2004; Pizer et al., 1998, 2003) for more in-depth discussions of the existing approaches.

The MAT can be directly extended to 3D as the centers of maximal inscribed spheres. For general 3D shapes, the MAT is then a 2D manifold.

Centerline-based vascular models 3D elongated structures such as vessels can be regarded as primarily 1D, with a prominent main axis. Theoretically, the MAT of a 3D vessel shape would indeed reduce to a 1D curve in some degenerate cases. This would require perfect symmetry conditions that are very rarely (if ever) met in practice. By *centerline-based* vascular models, we refer to models which impose the 1D topology of their core structure. More generally, centerline-based vascular segmentation methods can be viewed as techniques aiming at extracting the most significant, simplified and/or corrected subpart of the medial representation of a vascular network. Vascular centerlines correspond to curves *centered* inside the vessel (see Fig. 4.2), running along the main axis of the vessel. The eventual definition of *centered* varies with the model and practical extraction process. From an applicative point of view, centerline models are particularly interesting as they can be used directly to ease visualization, diagnosis and quantification tasks through the automatic generation of virtual navigation paths (fly-through) and curved planar reformation (CPR) views (see Fig. 7.26 in Chapter 7, for samples of CPR views).

Vascular centerline models classically incorporate smoothness assumptions. Such assumptions are either expressed implicitly, for instance through the extraction process (*e.g.* Kalman- or particle filter-based tracking, see Chapter 8), or explicitly. Among explicit centerline models, one can mention piece-wise linear curves in (Lacoste et al., 2006), *B-splines* in (Bouix et al., 2005; Flasque et al., 2001; Frangi et al., 2000; Larralde et al., 2003) and *cardinal splines* in (Wong and Chung, 2006).

Similarly to medial axes of the MAT, centerlines can be enriched with volume information, ensuring good reconstruction properties from a very compact model. The vessel surface can be conveniently defined by sweeping 2D cross-sectional contours along the centerline curve. Various cross-section geometric models have been used in the literature, from simple circles (see Fig. 4.2) and ellipses to star patterns (O'Donnell et al., 1998; Verdonck et al., 1995; Williams et al., 1997) and parametric curves (Li and Ourselin, 2003). The complexity of a cross-sectional model is directly linked to its descriptive power and potential accuracy. Additionally, the cross-sectional model significantly conditions (or is

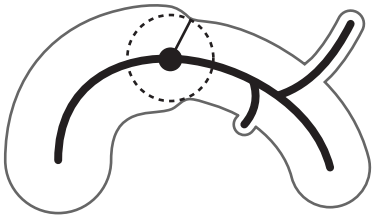


Figure 4.1: Medial Axis Transform of 2D shapes. The medial axes (bold black curves) correspond to the centers of maximal inscribed disks.

conditioned by) the choice of the vessel-dedicated features (see Sec. 3.2).

More generally, coupled centerline and cross-section models can be embedded in the larger family of *generalized cylinder* models which additionally incorporate techniques such as triangularization/meshing algorithms to maintain a coherent 3D surface (Ponce et al., 1989; Huang and Stockman, 1993; Terzopoulos et al., 1988; Zerroug and Nevatia, 1996). One of the most advanced models is the *extruded generalized cylinder* from (O'Donnell et al., 1994), which handles curved centerlines and varying cross-sections while avoiding torsion artifacts.

Some works alternatively use active surface evolution initialized from or constrained by a centerline curve (Frangi et al., 1999a, 2000; Montagnat, 1999; van Bemmel et al., 2003). The use of an existing centerline curve increases the robustness and efficiency of the extraction, while active contour evolution generally yields more accurate and more coherent results than independent optimizations of 2D cross-sections. The geometric model of the vessel surface can be implicit, embedded in the evolution scheme (*e.g.*, through curvature-based evolution), or explicit, as with the B-Spline surfaces from (Frangi et al., 1999a, 2000).

With regards to geometric modeling, it is generally assumed that there exist well-defined bifurcation points. Whereas the MAT properties ensure the existence of such points, it might not be the case for certain centerline-based methods. For instance, *core* extracted centerlines generally do not intersect in bifurcation areas, as thoroughly discussed in (Fridman, 2004). Disambiguation and correction schemes are then required to comply with a simple, coherent geometric model.

4.1.2 Description of our Model

Following the principles introduced in the previous section, the geometric model we propose can be considered as a coupled centerline- and cross-section-based model. Its main advantage is its particular discrete parameterization to constrain and reduce the parameter space. This parameterization is illustrated in Fig. 4.3.

Our model relies on a curved centerline structure and *circular* cross-sections, lying in locally orthogonal planes. The circularity assumption allows straightforward parameterization and was found reasonable for the description of coronary arteries, as depicted in Fig. 4.4. The centerline curve is discretized as a series of centerline points $\{p_t\}_{t=0,\dots,L}$, to which are associated radius values and tangential direction vector, noted $\{r_t\}$ and $\{d_t\}$, respectively. Radius values and tangent directions define cross-sectional contours. The most intuitive way to describe the model is then as a series of triplets $x_{0:L} = \{(p_t, r_t, d_t)\}_{t=0,\dots,L}$. The notation $x_{0:L}$ designates the vessel in its entirety, as a chain of elements $x_t = (p_t, r_t, d_t)$.

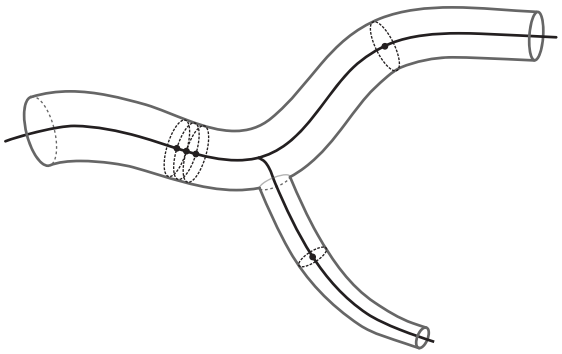


Figure 4.2: Coupled centerline and cross-section representation of 3D vessels. The centerline (bold black curve) corresponds to the curve centered in the vessel lumen. 2D cross-sectional contours are defined in planes perpendicular to the centerline (sweeping approach). Here, cross-sections are modeled as circles (black dashed circles).

In the remaining of the document, we refer to such atomic elements x_t as *states* of the model.

Note that we assume an order on the states, denoted by subscripts $t \in [\![0, L]\!]$. For coronary arteries, a natural ordering is from the ostium x_0 (origin of the artery branching off the aorta) to their distal ends x_L.

Tangential directions $\{d_t\}$ are defined thanks to *control* points $\{c_t\}$:

$$d_t = \frac{c_t - c_{t-1}}{\|c_t - c_{t-1}\|}$$

A key element of our model is to constrain control points with respect to centerline points:

$$p_t = \frac{c_t + c_{t-1}}{2} \tag{4.1}$$

In effect, this constraint links the positions of centerline and control points, indirectly coupling centerline points and tangent directions. Compared to simpler forward or backward schemes, our model maintains a stable definition of tangential directions even in areas of high curvature. This scheme is closely related to cardinal spline models, where tangents are defined as $d_t = a(p_{t+1} - p_{t-1})$, with a a tension parameter controlling the tightness of the curve. In this regard, our use of control points can be viewed as an artificial parameterization simplifying the formulation of our model and reducing its dimensionality. Control points constrain both the definition of tangential directions and the discretization of the centerline curve. States $x_t = (p_t, r_t, d_t)$ of our model can be described alternatively as $x_t = (c_{t-1}, c_t, r_t)$ given control points and radiuses. The overall dimensionality of our model is thus limited to 4D (3D control point locations + radiuses values).

By convention, we consider that the first centerline point p_0 is fixed and that the first tangential direction d_0 is defined solely by c_0 in a forward fashion. It is equivalent to considering an implicit control point $c_{-1} = 2p_0 - c_0$. Centerline points and corresponding tangents can be used to conveniently interpolate the centerline curve, *e.g.* using cubic Hermite splines. Finally, we will generally use cross-sectional radial directions (in the plane defined by d_i) as approximations of the local normal to the vessel surface (see Chapter 5). This approximation is generally reasonable given the slow scale variations of

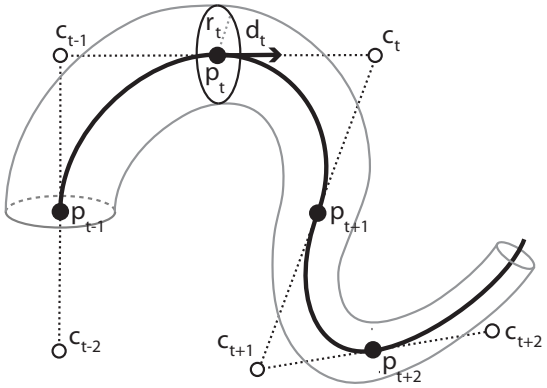


Figure 4.3: Discrete medial-based geometric model. The notation p_i designates centerline points with r_i associated radius value. The tangent direction d_t is defined as $d_t = \frac{c_t - c_{t-1}}{\|c_t - c_{t-1}\|}$ with c_t control points. Additionally, we set $p_t = \frac{c_t + c_{t-1}}{2}$.

coronary arteries (see Sec. 6.4.1), but may become inaccurate in presence of aneurysms and stenoses.

Our model can be straightforwardly extended so as to incorporate branchings, as illustrated in Fig. 4.5. To do so, one just needs to allow a control point c_t to have more than one successor. With respect to the centerline curve, this scheme has the advantage of directly defining branching points, at centerline points associated to multi-successor control points (p_t in Fig. 4.5).

One key parameter of our model is the spacing of successive control points $s = \|c_{t-1} - c_t\|$. This discretization step directly affects the expressive power of the model. As it gets smaller, the model is able to depict accurately even highly curved vessels. In this work, we use a fixed discretization step of the order of the data intra-slice resolution (0.3mm). Such a fine discretization allows an accurate capture of typical coronary arteries, as discussed in the next section.

4.1.3 Fitting the Model to Ground-Truth Delineations

The ground-truth delineations at our disposal consist of a series of centerline points, cross-sectional orientations and associated radius values. To match such unconstrained representations with realizations of our geometric model, we introduce a simple fitting procedure. Fitted models will be used in Sec. 6 to learn the different terms of our Bayesian model, built upon our geometric model. Additionally, the evaluation of the fitting errors confirm the adequacy of our geometric model in terms of expressive power.

Interpolation of ground-truth delineations The series of centerline points composing the ground-truth segmentations are relatively dense [1] but can be spaced unevenly as they originate from manual delineations. The first step was to continuously interpolate the ground-truth centerline (and associated radiuses) with cubic Hermite splines, making use of the available tangent definitions.

1. In general, ground-truth centerlines were extracted at a resolution of the order of the local radius.

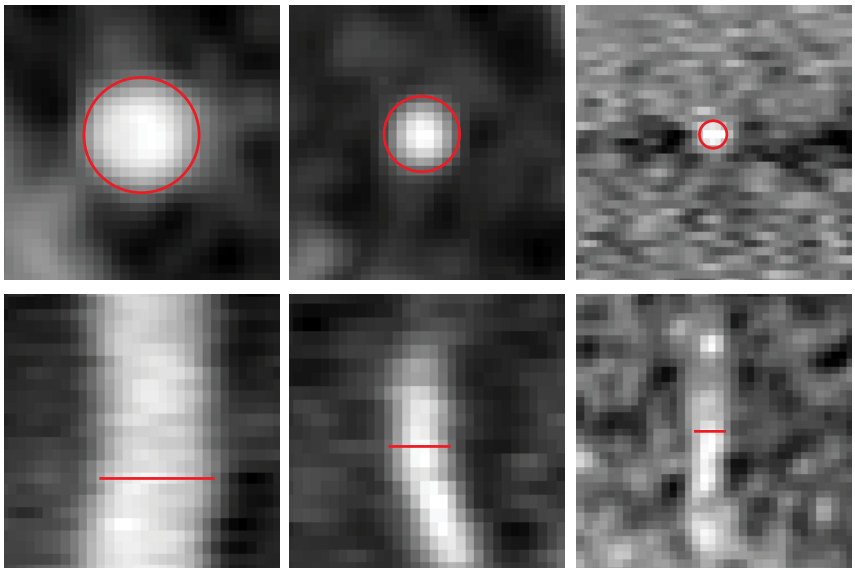


Figure 4.4: Examples of coronary cross-sections. Per column: sample of large- (radius $\simeq 3mm$), mid- (radius $\simeq 1.2mm$) and small-scale (radius $\simeq 0.6mm$) coronary arteries from real, unprocessed cardiac CTA data. Top row: cross-sectional view and circular approximation overlayed in red. Bottom row: associated transversal view.

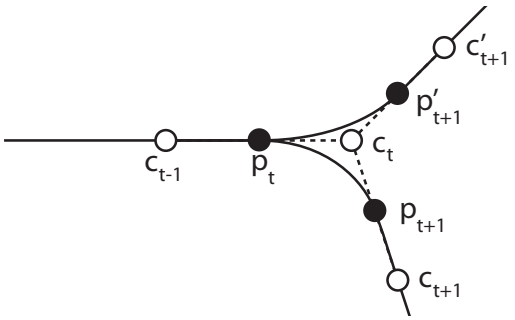


Figure 4.5: Geometric model: bifurcations. Branchings are handled by allowing a control point c_t to have more than one successor, here c_{t+1} and c'_{t+1}. The centerline point p_t is a branching point.

Iterative fitting procedure We use an iterative procedure to fit our geometric model to the interpolated ground-truth splines. This procedure is illustrated schematically in Fig. 4.6. The model discretization step s is chosen relatively small ($s = 0.3mm$ in our experiments), of the order of the intra-slice resolution of the data. The same resolution is used in the implementations of both the extraction schemes we propose (Chapters 7 and 8).

Given previous centerline and control points p_t and c_t, the next model centerline point p^*_{t+1} is selected along the ground-truth centerline as the first point, from p_t, at a distance $\frac{s}{2}$ from c_t. The corresponding control point c^*_{t+1} is given by $c^*_{t+1} = 2p^*_{t+1} - c_t$ from the model. The corresponding radius r^*_{t+1} is given by the ground-truth radius associated with point p^*_{t+1}.

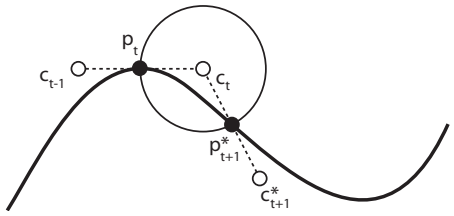


Figure 4.6: Fitting the geometric model to ground truth delineations (fixed discretization step s). See text for details.

	Root mean square error	Maximal error
Fitting step $s = 0.3mm$	0.078 ± 0.011 rad.	0.23 rad.

Table 4.1: Tangential approximation errors from the model fitting procedure. Results over the entire database (61 datasets).

One can note that the fitting process favors centerline localization over tangent direction. As stressed previously, our primary focus is on the centerline location, scale and direction being used to guide its segmentation. The fitting process introduces tangent approximation errors, given in Table 4.1. For a discretization step of $s = 0.3mm$, it yields a mean angular error of 0.078 radians (4.47 degrees) with respect to known ground-truth tangents. This error remains limited (maximal error of 0.23 rad. or 13 degrees), confirming the adequate expressive power of our geometric model.

A corollary question is whether our geometric model is able to capture the continuous centerline and radius variations of the ground-truth despite their discrete nature. We performed "reconstructions" of the fitted models through cubic Hermite spline interpolation and compared them to the ground-truth data. The errors with respect to tangential orientation, centerline localization and radius are given in Table 4.2. The low errors obtained were actually expected, as the model resolution is generally finer than the original, discrete manual delineations, and that both are interpolated in the same fashion. The main difference is that the interpolation of the models uses slightly erroneous tangents. This experiment confirms that tangential angular errors induce very limited centerline and radius deviations.

4.2 Appearance Model: Hounsfield Intensity Range

In CT image data, the intensity of a given voxel corresponds to its local (integrated) attenuation value (radio-density). Intensity values, also referred to as CT numbers, are expressed on a scale of arbitrary units called Hounsfield Units (H.U.). The origin of this scale is traditionally associated to the attenuation of water. This results in calibrated intensity

	Centerline error	Tangential error	Radius error
Fitting step $s = 0.3mm$	0.009mm (0.035)	0.075 rad. (0.23)	0.011mm (0.041)

Table 4.2: Reconstruction errors when interpolating fitted models and comparing it to the (interpolated) ground-truth data. Results over the entire database (61 datasets). Errors are given as root mean square errors, with the maximum error encountered in parentheses.

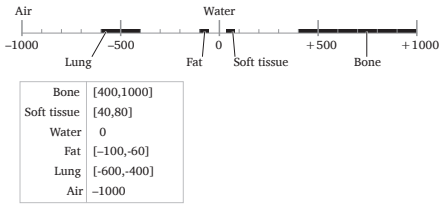


Figure 4.7: Hounsfield scale (expressed in CT numbers).

scales in which different tissues, of different attenuation coefficients, occupy different, predictable CT ranges (see Fig. 4.7). For our coronary application, we can focus our analysis to a limited intensity range corresponding contrast-enhanced blood. As a pre-processing step applied to the original data, we perform the following intensity transformation:

$$\tilde{I}(x) = \begin{cases} T_{low} & \text{if } I(x) < T_{low} \\ T_{high} & \text{if } I(x) > T_{high} \\ I(x) & \text{otherwise} \end{cases} \tag{4.2}$$

where $I(x)$ is the original CT number of a voxel x and $\tilde{I}(x)$ its transformed intensity value. The lower and upper thresholds T_{low} and T_{high} restrict the intensity dynamics within the range $[T_{low}, T_{high}]$. We experimentally set $T_{low} = -24H.U.$ and $T_{high} = 576H.U.$, defining a conservative range comprising contrast-enhanced arterial blood. T_{low} is set higher than lung CT numbers and slightly lower than cardiac soft tissues (myocardium). The upper threshold T_{high} corresponds to a conservatively high value separating contrast-enhanced blood from high-density structures such as calcifications. Original intensities lying in the range $[T_{low}, T_{high}]$ potentially correspond to contrast-enhanced blood and are left unchanged.

An illustration of the practical effects of this intensity transformation is given in Fig. 4.8. First, it attenuates contours at the lung-heart interface by raising the intensity of lung tissues close to that of the myocardium. This limits potential false positive responses of our contour-, flux-based feature (Chapter 5) in these areas and thus eases the extraction of coronaries running along the heart surface. Second, the intensity of calcified areas is lowered to the upper threshold. We emphasize that conceptually, this translates in the *inclusion* of calcifications in the coronaries. Other works in the literature advocate the *exclusion* of calcifications (Friman et al., 2008b; Metz et al., 2008a) as they are not part of the lumen. Our choice is motivated by the observation that in cases of severe atherosclerotic calcified plaque, the visible lumen can become extremely thin, so that excluding calcifications may result in near complete destruction of the local image structure. Calcified plaque, lying inside the artery, is actually an important structural clue we did not wish to suppress at such a low-level pre-processing stage. In our tests, including calcifications proved to be highly beneficial to the robustness of our extraction schemes in severely calcified cases. A counterpart of this choice is that the extracted centerlines may not always correspond to true lumen centerlines. A subsequent local correction step could be employed to exclude calcified areas *a posteriori* and recenter the centerline in the lumen.

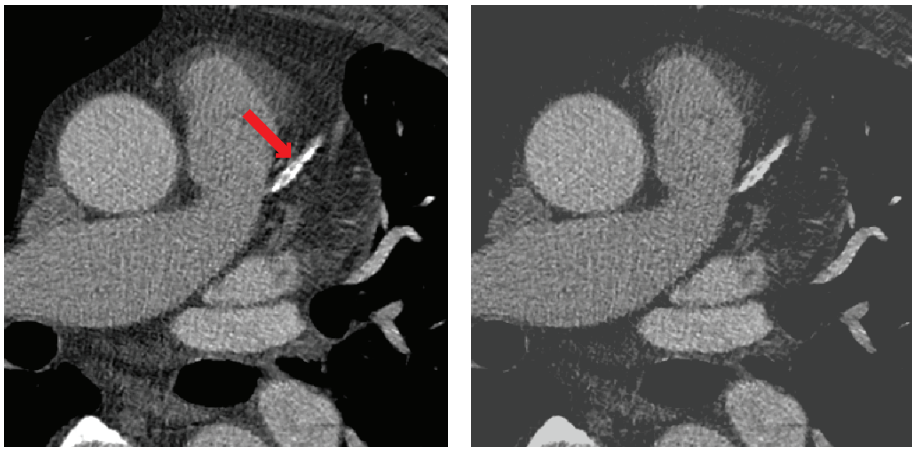


Figure 4.8: Illustration of our intensity mapping pre-processing step. Left: original data, with severe atherosclerotic calcified plaque. Right: intensity-transformed data.

4.3 Radial Appearance: Bar-Convolved Model

In Sec. 3.1.3 of our bibliographic review, we discussed existing *radial* appearance models to describe intensity variations along cross-sectional profiles. In our work, we use the bar-convolved profile (Krissian et al., 2000; Wörz and Rohr, 2004, 2007, 2008) to generate artificial test cases in Sec. 5.4. Although our approach does not explicitly rely on this model, it conditions the relevance of some of our experiments and discussions.

One of the most common radial models is the *Gaussian* profil, which models a vessel as a 3D Gaussian line, with higher intensities along the centerline and Gaussian attenuation toward the background (Friman et al., 2008a; Gong et al., 2003; Krissian et al., 2000; Rohr and Wörz, 2006; Shikata et al., 2004). It is particularly appealing from a mathematical point of view, as it allows various analytical derivations, from estimation bounds (Rohr and Wörz, 2006), to Gaussian scale-space analysis (Krissian et al., 2000). Such a model is relatively well suited to small-scale blood vessels in CTA acquisitions (Friman et al., 2008a). On the other hand, it does not faithfully model larger vessels, which tend to exhibit a plateau-like profile. The *bar-convolved* profile, which consists in a plateau convolved with a Gaussian kernel, accounts for this observation (Krissian et al., 2000; Wörz and Rohr, 2004, 2007, 2008). The homogeneous plateau models the ideal vessel while the Gaussian convolution accounts for PSF-related spatial blur. Model parameters include contrast (height of the plateau), radius (width of the plateau) and Gaussian standard deviation σ_m. As depicted in Fig. 4.9, the bar-convolved model faithfully reflects the appearance of coronary arteries of various sizes in our cardiac CTA data. Smaller arteries show a Gaussian-like profile and bigger ones exhibit a plateau.

We performed model fitting on radial profiles sampled from our entire ground-truth database (61 datasets, 858 delineated coronary branches) to estimate the average Gaussian standard deviation $\bar{\sigma}_m$ for this model. We obtained $\bar{\sigma}_m = 1.1$ voxels and reuse that value for the experiments in Sec. 5.4. Loosely speaking, the average fitted standard deviation $\bar{\sigma}_m$ can be viewed as an approximate, indirect measure of the average PSF on our database. A direct and accurate measure of the PSF for a specific device would require phantom acquisitions.

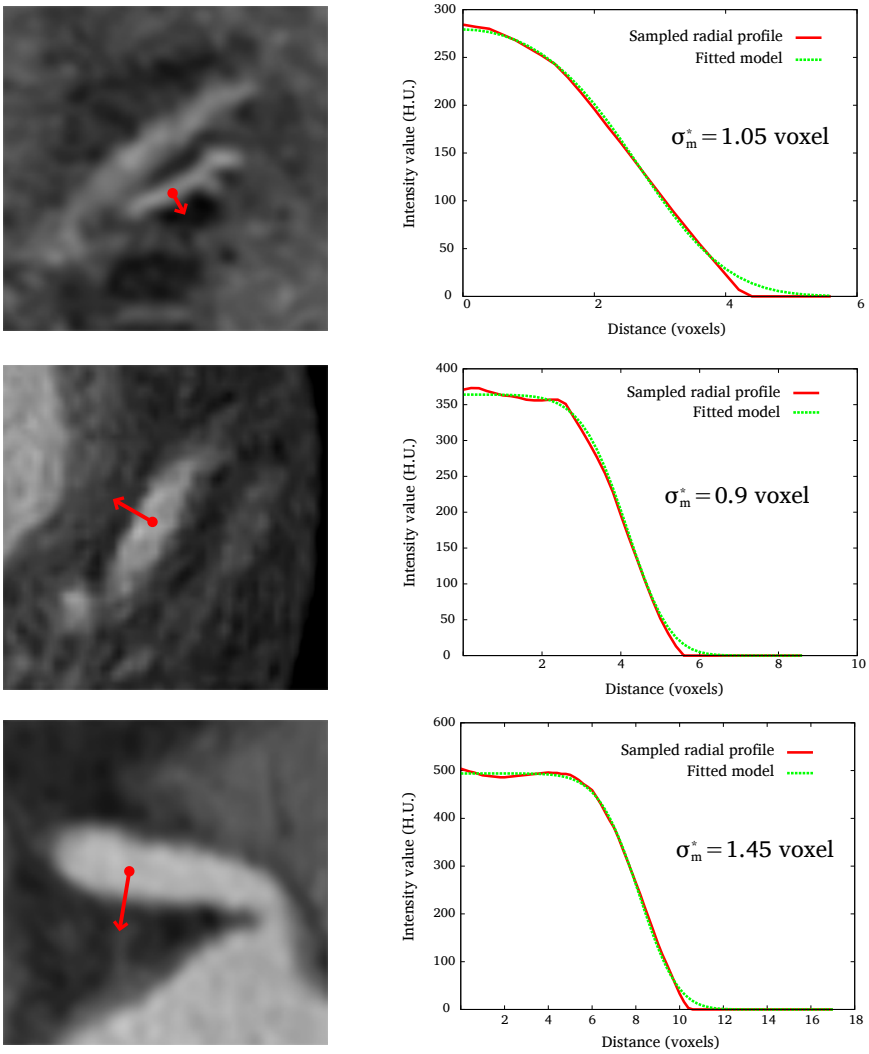


Figure 4.9: Bar-like convolved model and CTA coronary arteries. Fits of the bar-convolved model (Wörz and Rohr, 2008) on profiles sampled from a small (first row), medium (second row) and large (third row) coronary artery. Red arrows on multiplanar reformation views of the original data indicate the position and direction of the sampled profile. Graphs on the right depict the sampled intensity profiles overlayed with the fitted bar-convolved model. Fitted parameters σ_m^* are given for each case.

4.4 Discussion

From existing coupled centerline- and cross-section-based models, our geometric model inherits its compactness and simplicity. Its main particularity is the definition of *control* points to couple centerline positions and local tangential directions. This specific parameterization effectively limits the model dimensionality to 4D. This geometric model is not coronary-specific and could potentially be used for any application dealing with tubular structures, to the condition that the assumption of circular cross-section holds adequately.

The expressive power of our model depends principally on the discretization step s,

controlling the spacing of successive control points. For relatively small values of s, we showed that our model can accurately encode typical variations of coronary arteries. In fact, our model introduces only loose geometric constraints. It allows for a broad spectrum of possible centerline and radius variations, most of which are unlikely to correspond to coronaries. Further constraints on such variations are embedded as prior knowledge in the recursive Bayesian model discussed in Chapter 6.

We additionally discussed the exploitation of prior knowledge on the appearance of coronary arteries in cardiac CTA. We proposed a simple pre-processing step to focus on the typical intensity range of contrast-enhanced blood while limiting the influence of hypo-intense (lungs) and hyper-intense structures (calcifications). We finally illustrated the adequacy of the bar-convolved radial model, used to generate synthetic test cases for the experiments of the next chapter.

Chapter 5

Flux-Based Vessel-Dedicated Features: `Flux` and `MFlux`

This chapter is an updated and extended version of (Lesage et al., 2009c), published in IEEE International Symposium on Biomedical Imaging (ISBI 2009).

In this chapter, we discuss the image-based feature used for the evaluation of our geometric model, locally defined by a central position p, a radius r and an orientation d (Chapter 4). Our goal is the design of a single robust, discriminative and computationally efficient feature. We first present our seminal formulation of `Flux` , an oriented, local flux-based medialness measure. We then introduce a modification of that feature for increased discriminative power, yielding the `MFlux` feature. We discuss the links of `Flux` and `MFlux` with existing works and propose an efficient implementation. Through a series of experiments on synthetic and real data, we analyze the theoretical and practical properties of our features.

5.1 Description of `Flux` and `MFlux`

Gradient magnitude is a classical image measure used for contour detection. Besides magnitude, gradient flux-based segmentation methods exploit the *orientation* of gradient vectors by computing the gradient *flux* through the surface of the extracted object. As

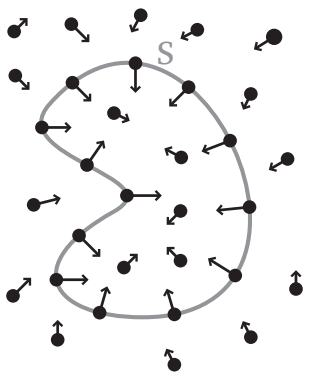


Figure 5.1: Gradient flux and surface alignment. The inward gradient flux is maximized when the object's interface S is aligned with the gradient vector field (arrows).

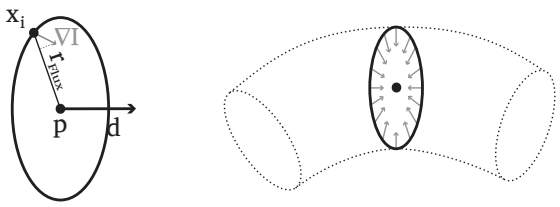


Figure 5.2: `Flux` feature. Left: the `Flux` response is computed over a discretized cross-sectional pattern defined by parameters (p, r_{flux}, d), with p the central point, d the orientation and $r_{\texttt{Flux}}$ the *test* radius (see text). At each cross-sectional test point x_i, the gradient vector $\nabla I(x_i)$ is projected on the inward radial direction, $u(x_i) = \frac{p-x_i}{\|p-x_i\|}$. Right: `Flux` response is maximized when aligned with hyper-intense tubular image structures such as vessel, so that all image gradients over the cross-section test pattern have high magnitudes and point towards the cross-section's center p.

depicted in Fig. 5.1, the core idea of such approaches is to align the interface of the object with the underlying gradient vector field. For hyper-intense structures such as our vessels of interest, this is done by maximizing the *inward* gradient flux through the boundary [1]. The flux $F(S)$ through a surface S is given by:

$$F(S) = \int \langle \nabla I, n \rangle ds \tag{5.1}$$

where ds is an infinitesimal surface patch, $\langle \nabla I, n \rangle$ denotes the scalar product (projection) of the gradient vector ∇I on the unit *inward* surface normal n.

To the author's knowledge, flux maximization for vascular segmentation was first introduced in (Vasilevskiy and Siddiqi, 2002) and later extended and combined within other frameworks (Descoteaux et al., 2004; Kimmel and Bruckstein, 2003; Holtzman-Gazit et al., 2006). In (Vasilevskiy and Siddiqi, 2002), the authors highlight important connections with edge detection techniques, mean curvature evolution, *gradient vector flow* formulations from (Xu and Prince, 1998) and non-linear diffusion schemes.

As demonstrated in (Vasilevskiy and Siddiqi, 2002) with a level-set implementation, flux-based approaches are well suited to the extraction of thin, low-contrast vessels. In the context of general curve evolution, Vasilevskiy and Siddiqi (2002) approximate flux measures through local multiscale spheres. An efficient implementation for such spherical flux can be found in (Law and Chung, 2009). For our simple, parametric model, we can alleviate that approximation by considering circular cross-sections as local surface patches. This observation leads to the formulation of our first feature, `Flux` .

5.1.1 `Flux`

Local parameters (p, r, d) of our geometric model define a cross-sectional contour $C(p, r, d)$. The `Flux` feature measures the local flux contribution along a cross-sectional pattern $C(p, r_{\texttt{Flux}}, d)$, with a *test* radius $r_{\texttt{Flux}}$ closely related, but not necessarily exactly equal to the *model* radius r. As discussed in Section 5.4.3, the relation between *test* and *model* radiuses $r_{\texttt{Flux}}$ and r depends on the radial appearance model considered.

1. Formulas are straightforwardly inverted for hypo-intense objects.

After equi-angular discretization of the cross-sectional contour $C(p, r_{\texttt{Flux}}, d)$ into N points $\{x_i\}_{i=1,\dots,N}$ with $\|x_i - p\| = r_{\texttt{Flux}}$, we obtain the formulation of `Flux` :

$$\texttt{Flux}(p, r_{\texttt{Flux}}, d) = \frac{1}{N} \sum_{i=1}^{N} \langle \nabla I(x_i), u_i \rangle \tag{5.2}$$

where $u_i = \frac{p - x_i}{\|p - x_i\|}$ is the inward unit radial direction (see Fig. 5.2). As vessels generally exhibit slow radius variations, radial directions u_i give a reasonable approximation of the local normals to the surface. `Flux` can be viewed as a discretized, approximated measure of the general flux formulation (Eq. 5.1) making use of the specificities of our geometric model.

5.1.2 `MFlux`

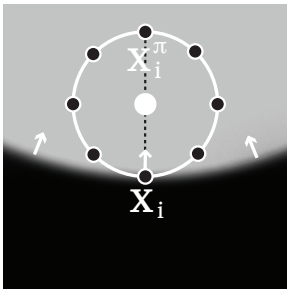


Figure 5.3: Step-edge response issue and `MFlux` test pattern. A linear feature such as `Flux` exhibits unwantedly high response for small test radiuses at the surface of large scale structures. Such false positive situations are characterized by asymmetric gradient flux contribution over the tested cross-sectional contour. `MFlux` considers only the minimal contribution between diametrically opposed pair points (x_i, x_i^π) to alleviate this issue.

As will be illustrated in the subsequent experiments (see Fig. 5.19), a classical shortcoming of linear measures, including `Flux` , is their sensitivity to isolated *step edges* (Koller et al., 1995). In particular, `Flux` produces relatively high responses along the surface of large, hyper-intense structures such as heart chambers in cardiac CTA data. Similarly, if the vessel of interest is sufficiently large with respect to the image resolution, spuriously high responses are recorded at small scales, in the vicinity of the vessel wall (see Fig. 5.13 for an experimental illustration). As depicted in Figure 5.3, step-edge effects are characterized by *asymmetric* contributions to the cross-sectional flux; only cross-sectional points along the boundary of the large scale structure gather high gradient projections. Following the seminal idea of non-linear combination from (Koller et al., 1995) in 2D, we propose to pair diametrically opposed points (x_i, x_i^π) and retain the *minimal* flux contribution per pair. We obtain our second feature:

$$\texttt{MFlux}(p, r_{\texttt{MFlux}}, d) = \frac{2}{N} \sum_{i=1}^{\frac{N}{2}} \min(\langle \nabla I(x_i), u_i \rangle, \langle \nabla I(x_i^\pi), u_i^\pi \rangle)$$

with $x_i^\pi = x_{\frac{N}{2}+i}$ for an even number N of cross-sectional points. In (Koller et al., 1995), the geometric mean is proposed as a more tolerant alternative to the `min` operator. A

similar notion of penalizing asymmetric contributions was also used in (Pock et al., 2005) to derive a vascular-medialness feature, following the principles introduced in (Xu and Pycock, 1999) with the *Concordance Medial Axis Transform* (CMAT). Their idea was to weight each point contribution inversely to its deviation from the mean contribution, from the rationale that asymmetric contributions engender large variances in the distribution of the N contributions. Our approach is simpler, thus computationally more efficient, and parameter-free. It is also stricter, as symmetry constraints are enforced explicitly.

The 3D extension of (Koller et al., 1995) tests only two cross-sectional pairs and again combines responses with a `min`. In contrast, we test twice as many points (4 pairs, for $N = 8$ points, see Sec. 5.3) and find *averaging* pair responses to be a more adequate, less strict trade-off between detection robustness and suppression of step-edge responses. The advantages of `MFlux` over `Flux` , especially in terms of discriminative power, are discussed in the experiments of Section 5.4.

5.2 Related Work

Oriented flux Among existing flux-based features, we already mentioned the spherical isotropic flux measure used in (Vasilevskiy and Siddiqi, 2002), for which a particularly efficient implementation can be found in (Law and Chung, 2009). From their *oriented* formulation, `Flux` and `MFlux` are related to the *optimally oriented flux* feature from (Law and Chung, 2008), which can be seen as a flux-based anisotropy descriptor. The *optimally oriented flux* from (Law and Chung, 2008) aims at locally determining the principal flux-maximizing orientation, formulated as a generalized eigenvalue problem applied to local multiscale spherical regions. The frequency domain implementation of (Law and Chung, 2008) is adapted to the filtering of whole volumes, using a limited of set of pre-selected radiuses. On the other hand, `Flux` and `MFlux` are designed for a slightly different applicative scope, that is to explicitly test given parameters (position, radius, orientation) and sparsely explore the image data, in vessel tracking setups.

Seminal works on medialness features By design, `Flux` and `MFlux` are similar to existing cross-section-based features, sharing the seminal ideas from (Koller et al., 1995), considered as an early, 2D thus mono-ray approach that inspired many works on oriented medialness features (Fridman, 2004; Krissian et al., 2000; Pizer et al., 1998; Wink et al., 2000a; Gulsun and Tek, 2008b). The 2D steerable filters of (Koller et al., 1995) are based on differences of oriented Gaussian derivatives, computed at contour locations, on each side of an elongated structure. The use of Gaussian derivatives can be interpreted as a *regularized* computation of gradients, as discussed in the next paragraphs. The difference between the two shifted responses can be viewed as the sign change due to opposite inward radial directions. We also credit the initial idea of non-linear combination for step-edge response reduction, used in `MFlux` , to (Koller et al., 1995). As was discussed in the previous section, similar ideas were proposed independently in (Xu and Pycock, 1999) with the CMAT approach in the general context of medial axis extraction, and later exploited in (Pock et al., 2005) to derive a vascular-dedicated feature. An additional early effort worth mentioning is the work of (Poli and Valli, 1996), who focused on the computational efficiency (separability) of oriented filters based on shifted Gaussian derivatives. They additionally proposed a so-called *validation* scheme, relying on intensity criteria, to reduce isolated step-edge responses.

Different approaches to gradient regularization For flux computation, it is preferable to pre-regularize the gradient field for increased robustness. We integrate regularization by isotropic Gaussian smoothing in the numerical scheme for gradient field computation. With regards to flux computation, regularization is considered as preprocessing of the gradient vector field and is not necessarily performed by Gaussian filtering. For instance, it might be worth experimenting with anisotropic image regularization before gradient computation for `Flux` and `MFlux` . Gaussian-based regularization was chosen for its simplicity and computational efficiency, as discussed in Section 5.3. Such a classical scheme for gradient computation is shared by a number of works on vessel-dedicated features. In (Krissian et al., 2000; Wink et al., 2000a; Pock et al., 2005), Gaussian-smoothed gradient vectors are projected on radial directions of a cross-sectional test pattern, similarly to our approach. Krissian et al. (2000); Pock et al. (2005) employ circular cross-sectional patterns. In particular, Krissian et al. (2000) thoroughly study the theoretical and practical properties of their filter. Wink et al. (2000a) prefer a ray-casting threshold-based approach to extract vessel walls. Such a scheme is better suited for large scale vessels for which the circular model does not hold well, *e.g.* the descending aorta in abdominal CTA.

The `Core` technique from (Fridman, 2004) also relies on a circular gradient-based pattern with equi-angular rays. Instead of projecting Gaussian smoothed gradient vectors onto radial directions, their feature employs Gaussian derivatives computed with *directional* kernels oriented along the rays. One can note that both approaches are theoretically equivalent from a continuous point of view, due to the linearity and separability of the Gaussian. In fact, the main difference between `Flux` and the features from (Krissian et al., 2000; Fridman, 2004) lies in the use of regularization. In (Koller et al., 1995; Krissian et al., 2000; Wink et al., 2000a; Fridman, 2004; Pock et al., 2005), the Gaussian standard deviation varies with the targeted vessel radius, following principles of Gaussian linear scale-space theory (Lindeberg, 1994). More generally, the choice of the regularization derives from a trade-off between *detection* robustness and *localization* accuracy/discriminative power, as thoroughly discussed in (Fridman, 2004). Stronger regularization induces better detection, but lower discriminative power, as illustrated in the experiments of Sec. 5.4. Varying regularization with scale is generally a sensible trade-off between these two conflicting issues. It is particularly attractive for applications with high noise levels, such as ultrasound imaging, or applications dealing with large-scale vessels and requiring some tolerance to deviations from circular cross-sections, such as CT abdominal aneurysms. For our particular application, CT coronaries, good robustness levels can be reached with fixed, limited regularization, without sacrificing accuracy (Gulsun and Tek, 2008b). Also, the use of large Gaussian standard deviations makes the filter response sensitive to disturbance by nearby structures, as the derivative kernel may extend well beyond the boundary of the target object (Law and Chung, 2008). It finally leads to important computational benefits, as a single gradient field is needed and can be pre-computed, as discussed thereafter.

Classification of medialness features Following the classification of medialness measures from (Pizer et al., 1998), `Flux` and `MFlux` can be characterized as *offset* medialness features, exploiting *boundary* cross-sectional information, as opposed to *central* information at point p. Pizer et al. (1998) also state the difference between *linear* and *adaptive* medialness measures. By *adaptive*, Pizer et al. (1998) characterize techniques making use of data-dependent information, *e.g.* the eigenvalues and eigenvectors of the Hessian matrix, to orient the feature. The `Vesselness` feature from (Frangi et al., 1998) and the feature from (Krissian et al., 2000) are, respectively, examples of *central* and *offset* adap-

tive medialness features. Our approach is slightly different, as we do not orient our feature in function of the data and let the extraction schemes choose explicitly which orientations (and radiuses and positions) to test. We do not rely on local orientation estimations from the data and thus free ourselves from issues due to the potential unreliability of such estimations.

5.3 Implementation Details

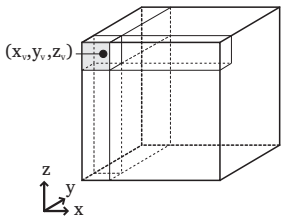


Figure 5.4: Block volume data structure. The original volume is subdivided into smaller blocks. Each block is encoded, in memory, as consecutive values following a classical lexicographic order. From original data coordinates (x_v, y_v, z_v), the block volume is addressed by first computing the coordinates of the block, then the coordinates of the data within the block.

The straightforward, naive implementation of `Flux` and `MFlux` can be optimized in a number of ways for increased computational efficiency.

Number of cross-sectional points In our implementation, we use $N = 8$ points, similarly to the `Core` feature from (Fridman, 2004). In our experiments, we observed that the use of less than 8 points significantly affected the robustness of the responses, especially for `MFlux` . Conversely, the gain in stability brought by using more than 8 points was, in our opinion, outweighed by the additional computational burden. Our choice is thus a trade-off between speed and robustness. A corollary advantage of using a fixed, low number of cross-sectional points is the possibility of pre-computing numerous quantities, such as the unit radial directions for a pre-defined set of orientations.

Pre-computation of the regularized gradient field We pre-compute the 3D gradient field on a fixed, isotropic regular grid. This involves three separate steps. First, we tri-linearly interpolate the original data to obtain an isotropic, resampled volume. Second, we map original intensity values to focus on a range corresponding to contrast-enhanced blood (see Sec. 4.2). Third, the gradient field is jointly computed and regularized using recursive implementations of Gaussian derivatives (Deriche, 1992). We use a standard deviation $\sigma_g = 0.3mm$, of the order of the image resolution, so as to regularize while preserving the smallest vessels. This pre-computation step, written in C++, takes on average 2 minutes for a 512^3 volume on a 2.16GHz `Intel Core Duo Intel` processor. The operations involved are adapted to straightforward parallelization. Using simple C++ `OpenMP` directives, we were able to almost double the computational speed. We expect similar proportional gains if more processors were available.

Memory- and cache-friendly organization of the gradient field Storing the 3D gradient field for an entire volume can be particularly memory intensive. We chose to quantize gradient vectors as 8-bit integer values (closest integer rounding), between -128 and 127 H.U. This yields a cost of 384MB (mega-bytes) for a 512^3 isotropically resampled gradient field. For our cardiac CTA application, 8 bits offer a value range sufficient to ensure acceptable accuracy of the measures [2]. Direct 32-bit floating point encoding would require 1536MB for a 512^3 gradient field. We tested the alternative of using 16-bit fixed point values for increased value range (768MB) and did not observe significant differences with 8-bit encoding. We eventually favored memory efficiency over accuracy given the typical limitations of today's workstations.

Closely related to the dimension of the gradient maps, the efficiency of memory access to the gradient values is one of the principal computational bottlenecks of our approach. Our extraction processes (Chapters 8 and 7) require the evaluation of thousands, potentially millions of evaluations of `MFlux` , yielding proportionally numerous memory accesses to retrieve pre-computed gradient values. Such disorganized, unpredictable accesses result in costly cache misses which dramatically deteriorate the computational efficiency. To alleviate this issue, we first organize the gradient field in a cache-friendly manner. First, instead of using 3 separate volumes to encode x-, y- and z-gradients, we encode the gradient field as a single volume with consecutive x,y and z gradient components. Second, we use a *block* volume data structure (see Fig. 5.4), dividing the original field into cubic sub-volumes (blocks). Each block is encoded as consecutive values in memory. The key idea is that a block regroups gradient vectors from a cubic portion of the 3D space, where classical encoding would follow a lexicographic line-row-depth order. As our exploration of the 3D space is done in a coherent fashion, memory accesses are likely to correspond to very localized space areas, thus to a low number of active blocks which can be kept in the processor's cache memory. We use cubic sub-volumes consisting of 32^3 gradient vectors, yielding a small memory footprint of 96KB per block. The use of such a data structure implies an additional cost due to its more complex, twofold addressing scheme, but it is greatly outweighed by its cache-friendliness. The practical gain depends on the extraction scheme and on the data at hand (in particular, on the spatial spread of the arterial tree). On average, we recorded a 30% speed gain over the naive implementation.

Gradient vector interpolation Gradient vectors at locations x_i are obtained by interpolation from the pre-computed field. We experimented with nearest neighbor and tri-linear interpolation. As an empirical trade-off between accuracy and efficiency, we use tri-linear interpolation for small test radiuses ($r_{\texttt{Flux}}$ or $r_{\texttt{MFlux}}$ inferior to 1.5mm) and nearest-neighbor interpolation for larger scales. From our experiments, the gain in accuracy using tri-linear interpolation was only significant for such small scales approaching the data resolution.

Practical computational efficiency The core computations in `Flux` and `MFlux` reduce to simple scalar products. Our C++ implementation makes use of vectorization instructions (`Intel SSE3`) for further performance gains. As a result, a single evaluation of `Flux` or `MFlux` takes on average 350 clock cycles, using nearest neighbor interpolation. In other terms, this allows us to compute more than 6 million evaluations per second on a 2.16GHz

2. For unit radial directions u_i encoded using floating point values and $N = 8$ points, the maximal error due to 8-bit encoding for the `Flux` and `MFlux` measures is 4 H.U.

processor. The tri-linear interpolation version is about 9 times slower, still resulting in almost 0.7 million evaluations per second. Such levels of computational performance are the keystone to the practical usability of the extraction schemes detailed in Chapters 8 and 7.

5.4 Experiments

This section presents a series of experiments on artificial and real 3D data exhibiting the practical and theoretical properties of the `Flux` and `MFlux` features, in comparison with a selection of features from the literature.

5.4.1 `Vesselness` , `Ribbon` and `Core` Features

5.4.1.1 `Vesselness` Feature

The first feature we use in our comparative study is the popular Hessian-based `Vesselness` measure from (Frangi et al., 1998).

Hessian-based multiscale features (Sec. 3.2.2) exploit the local geometric and radiometric information embedded in the Hessian matrix of the image data, through its eigenvalues and associated eigenvectors. For instance, in the case of tubular structures such as hyperintense vessels, the local direction can be estimated by the eigenvector associated with the lowest absolute eigenvalue, that is, the direction of lowest second order variations. In 3D, the two remaining eigenvectors form a basis for the cross-sectional plane. To cope with different vessel radii, the Hessian matrix is computed in a Gaussian scale space, *i.e.*, with second order derivatives of a Gaussian function with varying standard deviation.

A variety of different filters has been formulated following these principles (Frangi et al., 1998; Krissian et al., 2000; Li et al., 2003; Lin, 2003; Lorenz et al., 1997; Sato et al., 1998b; Shikata et al., 2004). They differ mainly in their use of the eigenvalues and eigenvectors of the Hessian matrix to discriminate tubular structures, *i.e.*, vessels, from other shapes. The Vesselness measure from (Frangi et al., 1998) was chosen for our comparitive evaluation for its popularity and its simple, intuitive formulation. We also implemented and tested the feature from (Sato et al., 1998b), which exhibited very similar behavior in our experiments.

We refer the reader to (Frangi et al., 1998) for the details of the `Vesselness` formulation. Throughout our experiments, we use the parameter values recommended in (Frangi et al., 1998).

Hessian-based features partly owe their popularity to their sound Gaussian multiscale framework. In vessel enhancement settings, their only explicit parameter is the image location to be tested. The maximum feature response is simply taken over scales, thanks to proper normalization of the Gaussian derivatives (Lindeberg, 1994). In our experimental setup, we are interested in the behavior of the `Vesselness` measure over scales, that is, its scale-space discriminative power. The scale (standard deviation of the Gaussian) is thus considered as a parameter of the feature that we refer to as $\sigma_{\texttt{Vesselness}}$.

5.4.1.2 `Ribbon` Measure

The second feature we use for comparison purposes is a simple cross-sectional region-based measure, the `Ribbon` measure from (Florin et al., 2005). It measures the difference

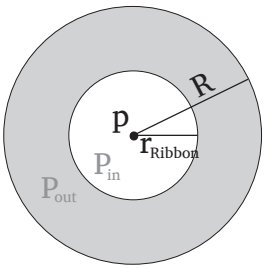


Figure 5.5: Sampling pattern for the `Ribbon` feature. See text for description.

between the mean intensity in a circular cross-section $C(p, r_{\texttt{Ribbon}}, \vec{d})$ and the mean intensity in an exterior ring, as depicted in Figure 5.5. the `Ribbon` measure is defined as:

$$\texttt{Ribbon}(p, r_{\texttt{Ribbon}}, d) = \frac{1}{N_{in}} \sum_{x \in P_{in}} f(x) - \frac{1}{N_{out}} \sum_{x \in P_{out}} f(x)$$

where $f(x)$ is the intensity for sample x and P_{in} and P_{out} are the set of sample points in the cross-section and in the exterior ring, respectively. The radius of the outer ring is set to $R = \sqrt{2} r_{ribbon}$ so as to have the same area and thus same number of samples inside (N_{in}) and outside (N_{out}). Sampling of P_{in} and P_{out} is performed at a subvoxelic resolution (typically 3 times the intra-slice resolution) using tri-linear interpolation. We restrict ourselves to circular cross-sectional areas, where (Florin et al., 2005) also considered ellipsoids.

This feature reflects the regional contrast between the cylinder and its local neighborhood. It takes high values for well-separated, low variance unimodal distributions P_{in} and P_{out}. In statistical terms, `Ribbon` is one-sided, *i.e.*, it introduces an assumption on the contrast polarity (hyper-intense vessels).

Naturally, more refined measures can be devised using region-based principles, *e.g.*, the measures from (Schaap et al., 2007a; Zambal et al., 2008) among many others. Our purpose here is to provide the reader with a simple region-based counterpart to the contour-based features `Flux` and `MFlux` .

5.4.1.3 `Core` Feature

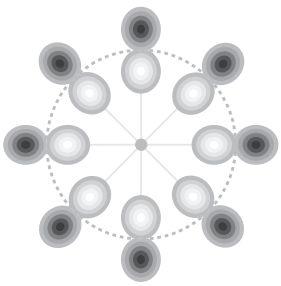


Figure 5.6: `Core` test pattern from (Fridman, 2004). Derivative Gaussian kernels are computed at the tips of equi-angular rays, oriented along radial directions.

The last feature in our study is the `Core` feature from (Fridman, 2004). As previously

discussed, the Core feature follows a cross-sectional, contour-based test pattern similar to Flux (see Fig. 5.6). In the Core approach, gradient information is retrieved by using Gaussian derivative kernels *oriented* along the rays. Additionally, the standard deviation of the Gaussian used for derivations is not fixed. Following the recommendations of (Fridman, 2004), it is set proportionally to the tested radius $r_{\texttt{Core}}$ as $\sigma_{\texttt{Core}} = \frac{r_{\texttt{Core}}}{4}$, with the restriction that $\sigma_{\texttt{Core}} \geq 1$ voxel.

5.4.2 Experimental Setup

For our experiments, artificial vessels are generated from straight tubes, following the bar-convolved appearance model from (Wörz and Rohr, 2007). This radial appearance model is faithful to the true appearance of vessels in our acquisitions. In particular, it better models large coronary arteries than a Gaussian radial intensity profile, as discussed in Sec. 3.1.3. Unless noted otherwise, we use a contrast of 300 H.U. between vessel and background and a standard deviation of σ_m =1.1 voxel for the bar-convolved model. These correspond to typical values of contrast and point-spread functions of our cardiac CT data (see Sec. 4.3).

Each feature was tested on a range of 10 radiuses between 1 and 10 voxels, in accordance with the typical resolution and range of radii in our data. Unless noted otherwise, 1002 test orientations (for Ribbon , Core , Flux and MFlux) were distributed pseudo-uniformly on a unit sphere, using a frequency-10 geodesic sphere layout (see Appendix A). This guarantees a relatively fine angular resolution (0.13 rad) ruling out most discretization artifacts from our analysis.

The features take values in different ranges. For comparison purposes, we scaled all responses between 0 and 1, 1 being the maximal response obtained over the parameters for the noiseless, maximum contrast bar-convolved straight cylinder. For Ribbon , Core , Flux and MFlux , negative responses were set to zero, as we are primarily interested in hyper-intense structures. For data such as cardiac CTA, negative responses happen only sparsely, for particular geometric configurations.

5.4.3 Scale Parameter *vs.* Theoretical Radius

All the features of our comparative study have scale parameters, related but not necessarily equal to the theoretical radius of the vessel. Vesselness has a Gaussian standard deviation parameter $\sigma_{\texttt{Vesselness}}$ controlling the Gaussian scale-space, Ribbon , Core , Flux and MFlux have test radiuses $r_{\texttt{Ribbon}}$, $r_{\texttt{Core}}$, $r_{\texttt{Flux}}$ and $r_{\texttt{MFlux}}$, respectively. For each feature, we derived the mapping between the theoretical radius and the optimally responding parameter value. This mapping is dependent on the feature at hand and on the radial appearance profile considered. For instance, in the case of Gaussian radial appearance profiles, Sato et al. (1998b) have shown, analytically, the relation $\sigma^* = \frac{r}{\sqrt{2}}$ for their filter, similar to the Vesselness measure. For the bar-convolved model, analytical derivations are much more intricate. We used numerical optimization, following the guidelines of (Krissian et al., 2000). For a given target radius r yielding a given bar-convolved profile, we optimized the scale parameters of the different features so as to maximize their respective responses.

Results of our numerical experiments are given in Fig. 5.7. For instance, it can be noticed that the optimal parameter $r_{\texttt{Ribbon}}$ is generally superior to the target radius. Of particular interest to us, it can be seen that optimal parameters $r_{\texttt{Core}}$, $r_{\texttt{Flux}}$ and $r_{\texttt{MFlux}}$

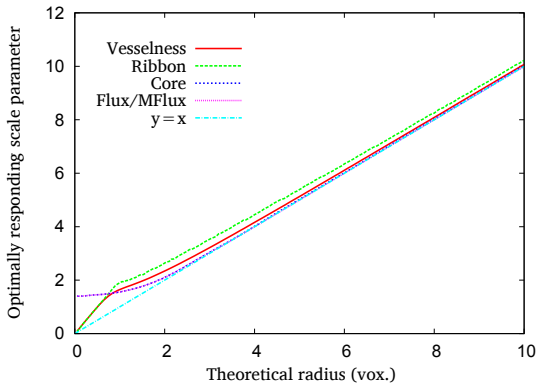


Figure 5.7: Relations between target vessel radius and optimally responding scale parameters $\sigma_{\texttt{Vesselness}}$, $r_{\texttt{Ribbon}}$, $r_{\texttt{Core}}$, $r_{\texttt{Flux}}$ and $r_{\texttt{MFlux}}$ for `Vesselness` , `Ribbon` , `Core` , `Flux` and `MFlux` features, respectively. Numerical optimization for a bar-convolved radial appearance model with $\sigma_m = 1.1$ voxel.

are almost exactly equal to the target radius for radiuses superior to 2 voxels. This is understandable as `Core` , `Flux` and `MFlux` are all gradient-, contour-based features. For large radiuses (with regards to σ_m), the inflection point of the radial profile, *i.e.*, the distance maximizing radial intensity gradient, corresponds to the theoretical radius r. More generally, the relations between optimal scale parameter and theoretical radius are mostly linear for large target radiuses, but differ slightly in small radius ranges. Thorough analysis of these results would require further, careful discussions with regards to each feature individually. In the context of our study, we use these mappings to ensure the fairness of the comparisons. For a given target radius, we select the optimal scale parameter for each feature.

5.4.4 Localization

The first experiment illustrates the sensitivity of the features to location eccentricity. Scale and orientation parameters are kept fixed to their optimal values. Results for different vessel radiuses are given in Figure 5.8. As expected, all feature responses are maximum at central positions and drop to zero about or slightly outside the vessel wall. More interesting are the differences of behaviors as a function of the target vessel radius. `Vesselness` shows a consistent, scale-independent behavior, a desirable property for its seminal purpose, vessel enhancement (Frangi et al., 1998). In contrast, the decrease rate of `Flux` mainly depends on the contour blur (σ_m for the bar-convolved model) and on the regularization of the gradient field (σ_g). As a result, `Flux` appears as increasingly discriminative relative to the vessel size. `MFlux` magnifies this advantage by penalizing asymmetric flux contributions arising with eccentricity.

On a side note, on can observe that `Core` and `Flux` behave almost identically for radiuses inferior up to 4 voxels. For such small radiuses, the `Core` feature employs a derivative standard deviation $\sigma_{\texttt{Core}} = 1$ voxel, equal to that of `Flux` . The differences observed are then due to the different numerical implementations (oriented Gaussian convolution for `Core` *versus* projection of regularized gradient vectors for `Flux`). For radiuses larger than 4 voxels, `Flux` becomes more discriminative than `Core` , due to the stronger

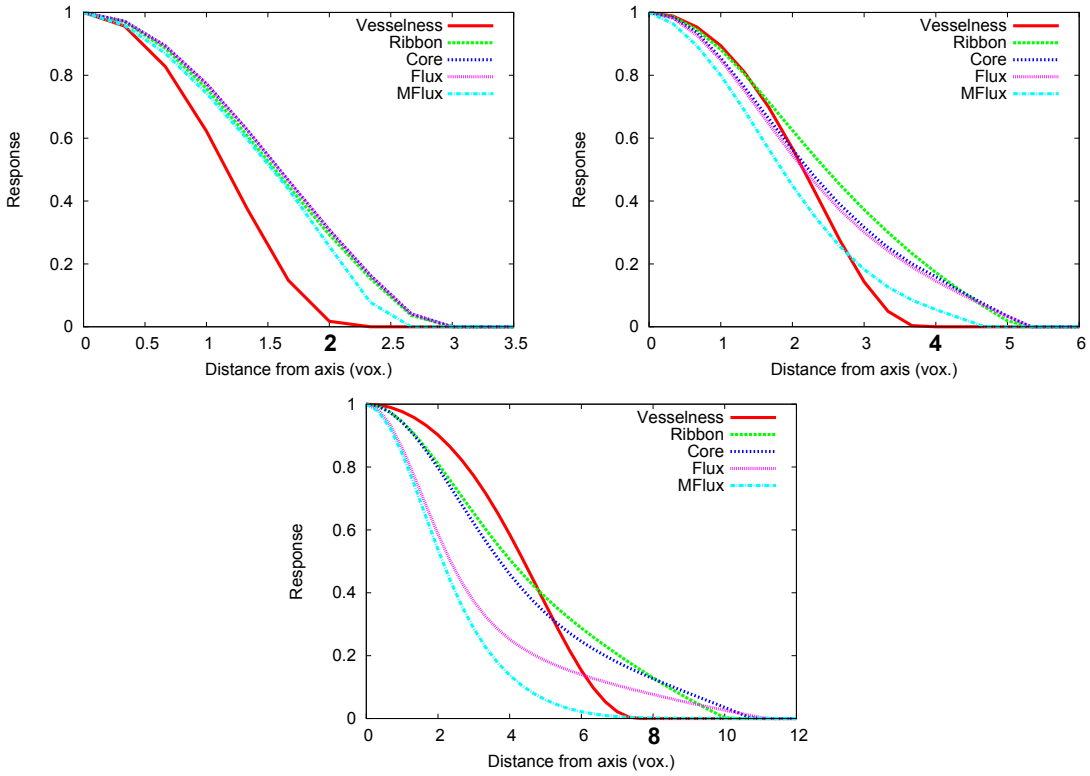


Figure 5.8: Responses over positions. From left to right: test for a bar-convolved straight cylinder of radius 2, 4 and 8 voxels. The response is given as a function of the distance from the axis. Scale and orientation are fixed to their optimal values.

regularization used by the latter.

5.4.5 Radius/Scale

Fig. 5.9 illustrates responses over target radiuses for central positions and fixed orientation. All features show a maximal response about the correct scale (4 voxels radius). In such symmetric situations, `Flux` and `MFlux` behave identically. Their responses are roughly symmetric about the correct radius. Because of their contour-based scheme and their scale-independent, limited regularization, they drop faster than the other features for tested radiuses larger than the true value. Additional illustrations of the scale-space behavior of the different features can be found in Section 5.4.7.

5.4.6 Orientation

We now illustrate the behavior of `Ribbon` , `Core` , `Flux` and `MFlux` with regards to orientation. `Vesselness` is left out of this experiment (and of all subsequent experiments involving orientation) as it does not provide us with explicit control of the orientation. We successfully checked that the Hessian eigenvector selected as principal direction coincides with the true vessel orientation. In this experiment, we set scale and position parameters to their optimal values. Because of the axial symmetry of the straight cylinder model, we

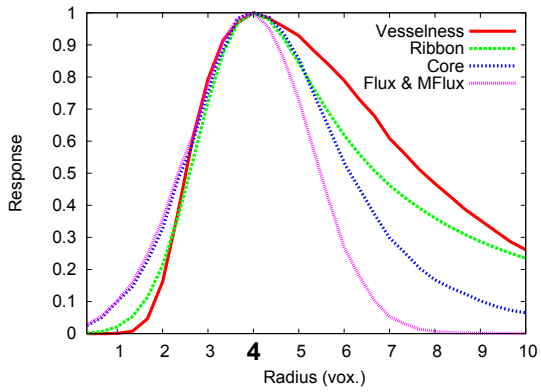


Figure 5.9: Responses over radiuses. Test for a bar-convolved straight cylinder of radius 4. The response is given as a function of the tested scale. Position and orientation are fixed to their optimal values (*i.e.*, only central positions and true vessel directions are tested).

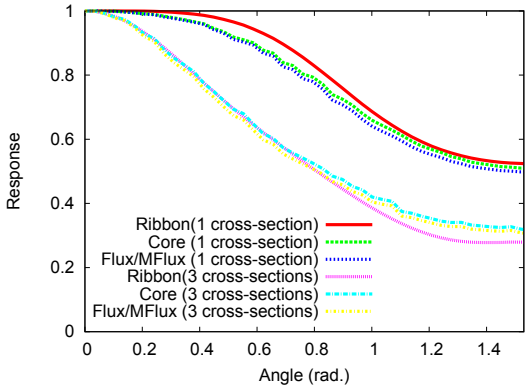


Figure 5.10: Responses over orientations. Test for a bar-convolved straight cylinder of radius 4. Responses are given as a function of the orientation (angle deviation from the true vessel direction). Position and scale are fixed to their optimal values. The "3 cross-sections" variants employ three successive cross-sections spaced by one voxel, along the tested orientation.

only test angular variations in a single plane, varying between 0 and $\pi/2$ radians about the true vessel orientation. Results are given in Figure 5.10.

`Flux` and `MFlux` seem marginally more sensitive to orientation changes than `Ribbon` or `Core` , although no feature shows a high discriminative power. The relatively low sensitivity to orientation changes is explainable as `Ribbon` , `Flux` and `MFlux` only evaluate a single cross-sectional plane. As depicted in Fig. 5.10, sensitivity to orientation can be increased by testing several successive cross-sections, *i.e.*, by exploiting the local elongation of the vessel. Following similar principles, many existing vessel-dedicated works employ oriented and *elongated* patterns, *e.g.*, (Schaap et al., 2007b; Wörz and Rohr, 2007; Friman et al., 2008a). Such designs implicitly rely on a local straight cylinder geometric model, which is stricter than our medial-based geometric model. As the tested pattern is lengthened, orientation sensitivity is increased, at the detriment of curved and scale-varying vessels.

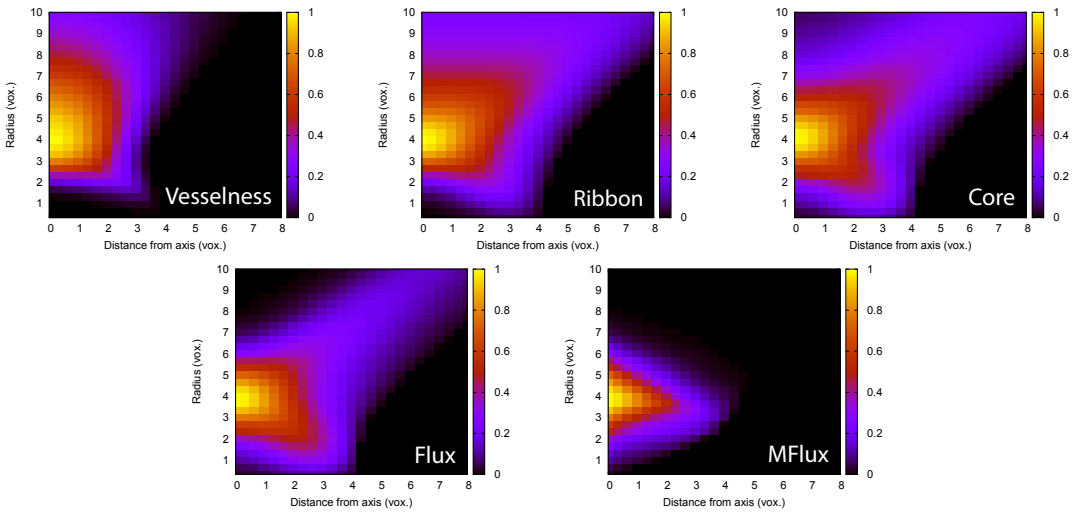


Figure 5.11: Joint variations of scale and position. Test for a bar-convolved straight cylinder of radius 4. Orientation is fixed to its optimal value. From top to bottom, left to right: `Vesselness` , `Ribbon` , `Flux` , `MFlux` .

We chose to limit our local feature to a single cross-section, knowing that successive cross-sections will be adaptively chosen and tested by the iterative extraction schemes we propose (Chapters 8 and 7), following our geometric (Chapter 4) and Bayesian model (Chapter 6).

Additional illustration of the behavior of the features with regards to orientation is given in Fig. 5.18 for an artificial branching network. For orientation $\frac{\pi}{2}$ for instance, it can be seen that the response for the small orthogonal branch is greatly attenuated for all features. Responses for the two Y-branching segments (at angles $\frac{\pi}{4}$ and $\frac{3\pi}{4}$) remain however relatively strong.

5.4.7 Scale and Position

Fig. 5.11 shows the response landscapes obtained when jointly varying the tested radius and the tested position. Of particular interest, this experiment illustrates the step-edge issue previously mentioned. It arises for small scales and eccentric positions and is particularly visible for `Core` and `Flux` . A similar effect is found for large scales and eccentric positions, more markedly for `Ribbon` and `Core` . The non-linear component introduced by `MFlux` dramatically reduces these issues, resulting in a much tighter response landscape. Step-edge effects (and their reduction by `MFlux`) are more clearly exhibited in Fig. 5.12 for a larger radius (10 voxels).

Fig. 5.13 depicts responses for straight vessels of different sizes, along with a vessel of slowly varying radius. It further illustrates that `MFlux` enjoys a strongly discriminative response with regards to eccentricity while considerably attenuating spurious step-edge responses.

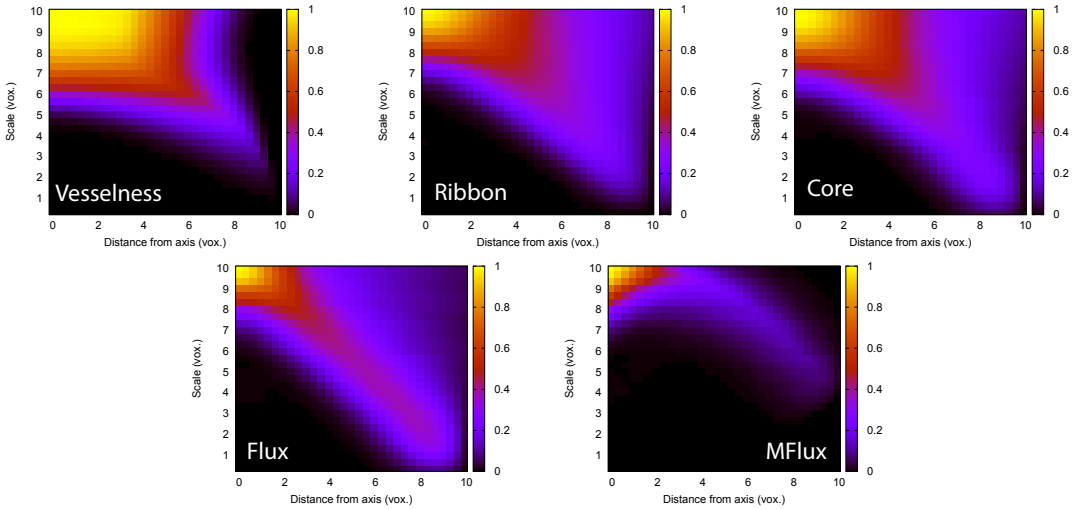


Figure 5.12: Joint variations of scale and position: illustration of the step-edge issue. Test for a bar-convolved straight cylinder of radius 10. Orientation is fixed to its optimal value. From top to bottom, left to right: `Vesselness` , `Ribbon` , `Flux` , `MFlux` .

5.4.8 Scale and Orientation

Joint variations of scale and orientation are depicted in Figure 5.14, for central positions of bar-convolved straight cylinder of radius 4 voxels. `Flux` and `MFlux` exhibit identical behavior as these test scenarios are symmetric. For all three features, it is confirmed that strongest response variations are found with scale changes, as was suggested by the experiment of Section 5.4.6. We also observe a slight correlated trend as scale and orientation increase. An erroneously large scale is compensated with increased angle deviation to still result in a relatively high response. This experiment principally illustrates the coupled discriminative advantage of `Flux` and `MFlux` over `Ribbon` and `Core` , both in terms of orientation and scale changes.

5.4.9 Orientation and Position

Joint variations of orientation and positions are depicted in Figure 5.15, for a bar-convolved straight cylinder of radius 4 voxels. Response variations seem relatively well uncorrelated, if one excepts the trend for positions about the vessel wall, for nearly orthogonal orientations. This phenomenon is especially visible for `MFlux` , as these situations have good symmetry properties. It is not particularly considered as an issue as it takes place far more the local maximum, at relatively low responses. We also expect real data, with non-perfectly tubular and curved vessels to be only marginally subject to such an effect.

Contour-based features `Core` , `Flux` and `MFlux` appear slightly more discriminative than the region-based, averaging `Ribbon` feature. `MFlux` 's response landscape is tighter than `Flux` 's and `Core` 's thanks to the penalization of flux asymmetry engendered by position eccentricity. Again, one can note that `Core` 's and `Flux` 's behaviors are almost identical.

5.4.10 Influence of Noise

Fig. 5.16 illustrates the impact of additive Gaussian noise (varying standard deviation σ_n on the responses of the different features. On the graphs of Fig. 5.17, we highlight two different aspects, namely the degradation of the response for optimal parameters (center of a vessel, optimal scale, optimal orientation) and spurious responses in the background.

As the strictest features of the panel, `Vesselness` and `MFlux` appear as the most sensitive to noise, in the sense that their responses for optimal parameters degrade faster with increasing noise. On the other hand, they are more robust to spurious background responses. More generally, the influence of noise is, as expected, greater for small scales. The response for thin vessels drops faster than for larger ones and average spurious background responses increase faster for small test radiuses. Such effects are implicitly taken into account in our likelihood model (Section 6.3). Overall, the behavior of all the features (of particular interest, `MFlux`) is deemed satisfactory given the typical noise levels affecting cardiac CTA data.

5.4.11 Behavior at Bifurcations

Fig. 5.18 illustrates the behavior of the different features for Y- and T-junctions. One can notice that the `Vesselness` feature is particularly affected by response drops as branching areas strongly deviate from the tubular shape prototype targeted by the Hessian eigenvalue ratios. Among remaining features, `MFlux` again exhibits the most discriminative response landscape, with a limited response drop for Y-junctions. The continuity of the maximum response ridge along the vessel centerline is only slightly perturbed in such symmetric situations. On the other hand, it is worth noting that, for all features, responses for the T-branching vessel (of smaller radius) appear as "disconnected" from the main tree. Logically, its corresponding branch only starts gathering high feature responses when it becomes distinct from the main vessel. In order to be able to retrieve such secondary branching segments, this suggests that the extraction scheme needs to be relatively tolerant to local response drops. This observation is kept in mind when designing stopping criteria for our extraction schemes in Chapters 8 and 7.

5.4.12 Influence of Surrounding Structures

In Fig. 5.19, we give response images for the test scenario of a straight vessel running nearby a large-scale hyper-intense structure (a heart chamber in cardiac CTA, for instance). The purpose of this experiment is twofold. First, it further illustrates issues with step-edge responses and their significant reduction with `MFlux` . Spurious response about the surface of the artificial heart chamber are particularly visible for `Ribbon` , `Core` and `Flux` . We emphasize that neither `Vesselness` nor `MFlux` completely suppress step-edge responses. Both features exhibit some very faint responses along the surface of the heart chamber, obtained for large test radiuses and explainable by the non-null surface curvature of the large-scale structure.

Second, this experiment illustrates the potential influence of neighboring structures on the response for actual vessels. All features show some deterioration of the centerline response when the vessel runs closest to the heart chamber. The `Vesselness` response is particularly sensitive to perturbations by nearby structures. `Vesselness` is a *central* medialness feature using second-order derivative kernels whose supports extend significantly outside the target object. To a much lesser extent, `Ribbon` and `Core` also show

response drops. `Flux` and `MFlux` , thanks to their contour-based formulation and their fixed, spatially limited regularization are least affected by closely adjacent structures.

5.4.13 Anatomical Anomalies: Stenoses and Aneurysms

In Fig. 5.20, we give response images for a straight tubular vessel affected by an artificial aneurysm (modeled as a larger sphere) and an artificial stenosis (double-cone). The `Vesselness` feature is, as expected, the most affected by such local geometric deviations from the tubular pattern. All remaining features exhibit similar behaviors, with drops of responses markedly visible at the start and end of the aneurysm (sharply widening and narrowing radius) and at the center of the stenosis. A key difference with `Vesselness` is that the cross-section-based features respond relatively highly in the aneurysm, which present first radius-varying but nevertheless circular cross-sections. A similar analysis can be drawn for the artificial stenosis. Cross-section-based features show a response gap narrower than for `Vesselness` . Responses are decreasing mostly due to the sharply narrowing then widening radius. Finally, it can be appreciated that `MFlux` has the most discriminative response landscape, while still showing relative robustness to such local anomalies.

5.4.14 Influence of Contrast

In Fig. 5.21 and 5.22, we illustrate the impact of varying contrast on the feature responses. The particular decrease rate of the `Vesselness` response is mainly explained by the "second order structureness" term controlling the `Vesselness` sensitivity to contrast:

$$S = (1 - \exp(-\frac{\|H\|_F^2}{2c^2}))$$

with H the Hessian matrix, $\|H\|_F$ its Frobenius norm and c a control parameter, set to half the value of the maximum Hessian norm in the image following advices from (Frangi et al., 1998).

The cross-section-based features, `Ribbon` , `Core` , `Flux` and `MFlux` are all directly dependent on region- of contour-based contrast from their formulations. This experiment confirms that their responses are exactly proportional to the actual (model) contrast. Unsurprisingly, lower contrast yields lower responses. On real cardiac CTA data, this raises issues for small arterial branches, which typically exhibit lower contrast mostly because of lower concentrations of contrast agent in distal parts of the coronary tree. Our likelihood observation model (Section 6.3), built upon statistical distributions of `MFlux` responses for real coronaries of various sizes, implicitly accounts for this typical loss of contrast.

5.4.15 Illustrations on Real Cardiac CT Data

In Fig. 5.23, we give sample response landscapes on volume of interest from real cardiac CTA datasets. For each voxel of the result images, the intensity encodes the maximum response over tested radiuses and orientations. This corresponds to a *projection* of the response landscape, equivalent to the filtering setup in which the `Vesselness` feature is classically used. We emphasize that we do not advocate the use of `Flux` or `MFlux` for enhancement filtering. The *optimally oriented flux* method from (Law and Chung, 2008)

would be more adapted for such setups. Our purpose here is simply to illustrate the selective behavior of the different features of the panel.

The `Vesselness` feature produces the smoothest results, as its Hessian-based, Gaussian scale-space design is well adapted to filtering applications. One can note nonetheless the drops of response at branchings and the weak response for smallest branches. For the remaining cross-sectional features, drops at branchings are more limited and smaller branches are better preserved. High responses are more tightly concentrated around the vessel centerlines, confirming the experiments on synthetic data and their increased potential for tracking tasks. Responses of `Ribbon` and `Core` appear smoother and less concentrated than those of `Flux` and `MFlux` . This is due, for `Ribbon` , to its region-based and averaging nature, and for `Core` , to its use of increased regularization for larger radiuses. This correlates with their generally lower sensitivity over positions (Fig. 5.8) and over tested radiuses (Fig. 5.9). The rows corresponding to multi-planar reformation views (MPR 2 and MPR 3) clearly exhibit spurious step-edge responses arising for `Ribbon` , `Core` and `Flux` at the surface of heart chambers. For `MFlux` , this effect is dramatically reduced while the response at true vessel locations is nearly unaffected.

In Fig. 5.24, 5.25 and 5.26, we give examples of `MFlux` responses on whole cardiac CTA datasets. One can note in particular the low responses in large scale hyper-intense structures such as the aorta and heart chambers, along the good capture of thin arterial branches. Despite its overall high discriminative power, `MFlux` still responds relatively strongly in some non-vascular areas, especially at elongated tips of cardiac ventricles, but also in the fibrous areas of the myocardium. Our Bayesian model (Chapter 6) and extraction schemes (Chapters 8 and 7) embed additional constraints on spatial connectivity and scale variations to increase robustness to such false positive responses.

5.5 Discussion

In this chapter, we presented an oriented multiscale feature, referred to as `Flux` , which measures the inward gradient flux through 2D circular cross-sections. From this seminal formulation, we derived a more discriminative variant, `MFlux` , aimed at reducing false positive responses in case of isolated step-edges. Through a series of experiments on artificial test cases, we studied the theoretical and practical properties of `Flux` and `MFlux` . In particular, we illustrated the gain in discriminative power brought by `MFlux` over `Flux` and other existing features from the literature.

We again emphasize that our purpose differs from vessel enhancement filtering. Enhancement setups classically aim at filtering the entire volume data, composing a single enhanced image from the aggregation of responses over a range of scales. To such aims, techniques such as (Frangi et al., 1998) or (Law and Chung, 2008) are particularly suited. Our focus was rather on the discriminative, online evaluation of sparse and specific realizations of our geometric model (specific positions, radiuses and orientations). In this context, the implementation of our feature is particularly straightforward and computationally efficient, making it well suited to the tracking extraction strategies discussed in Chapters 8 and 7.

One could consider using `MFlux` directly as a vessel detector. As illustrated in Fig. 5.24, 5.25 and 5.26, `MFlux` is generally able to well separate blood vessels from other hyper-intense structures. Isolated false positives still subsist, notably at elongated tips of cardiac ventricles. Additionally, `MFlux` responses vary with noise and contrast. In cardiac CTA data, thinnest arteries typically exhibit lower contrast and thus gather lower `MFlux`

responses. Such effects are taken into account in our approach, embedded in our Bayesian model (Chapter 6). In effect, we use `MFlux` responses *indirectly* to valuate the vessel and background likelihood terms of our model (Section 6.3).

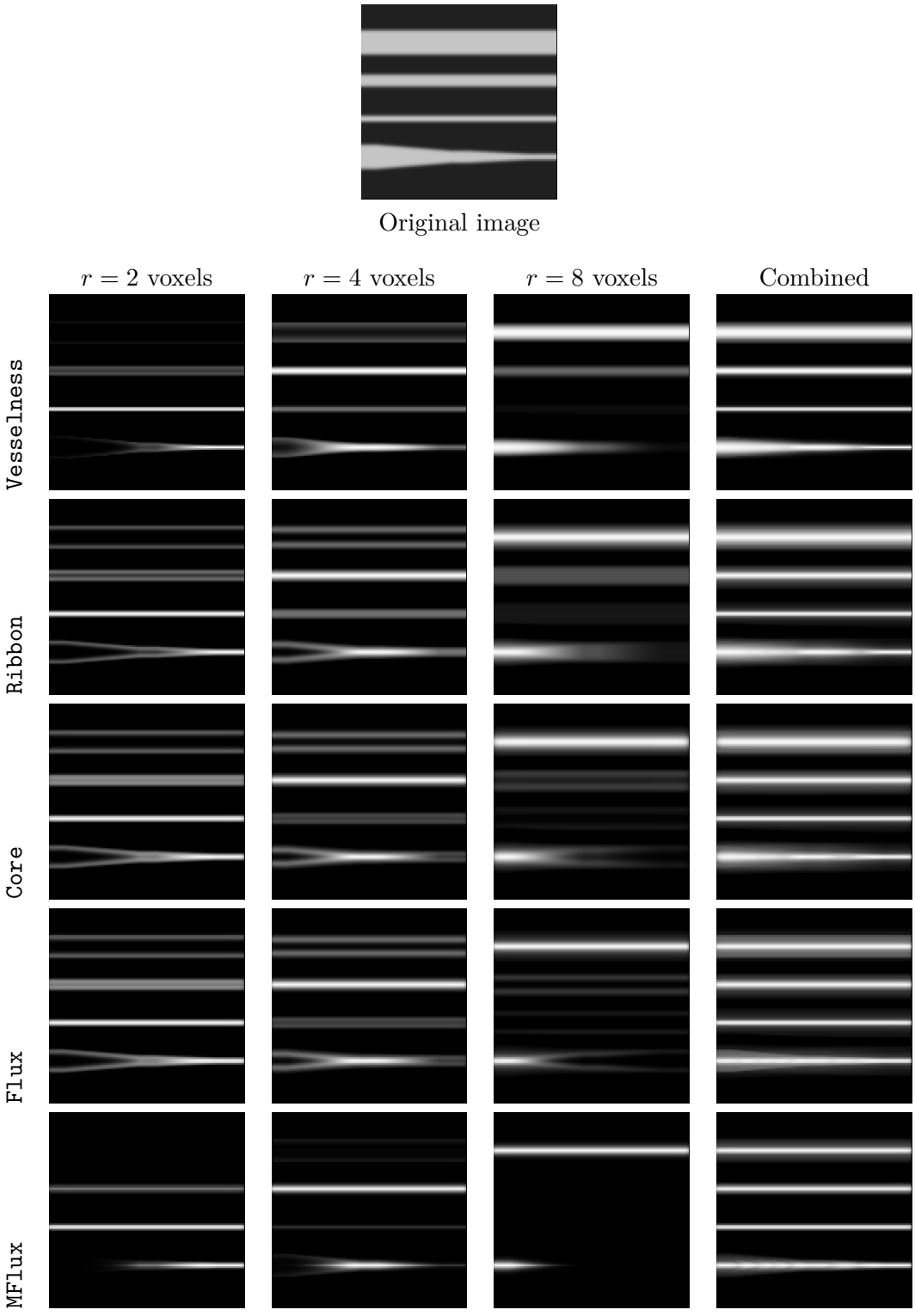


Figure 5.13: Responses for straight vessels of radiuses 2, 4 and 8 voxels and a vessel with varying radius (from 8 to 2 voxels radius). Each column gives the response, per voxel, at a target scale, for a single test orientation $(1, 0, 0)$. The last column gives the combined response (maximum response over tested radii).

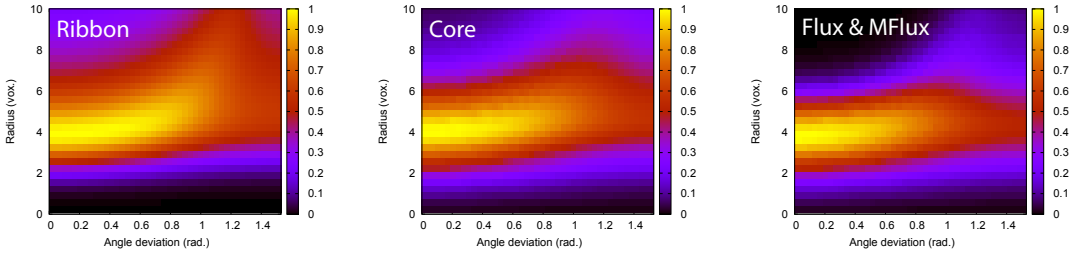


Figure 5.14: Joint variations of scale and orientation. Test for a bar-convolved straight cylinder of radius 4. Position is fixed to its optimal value (the center of the vessel). From left to right: `Ribbon` , `Core` , `Flux` and `MFlux` .

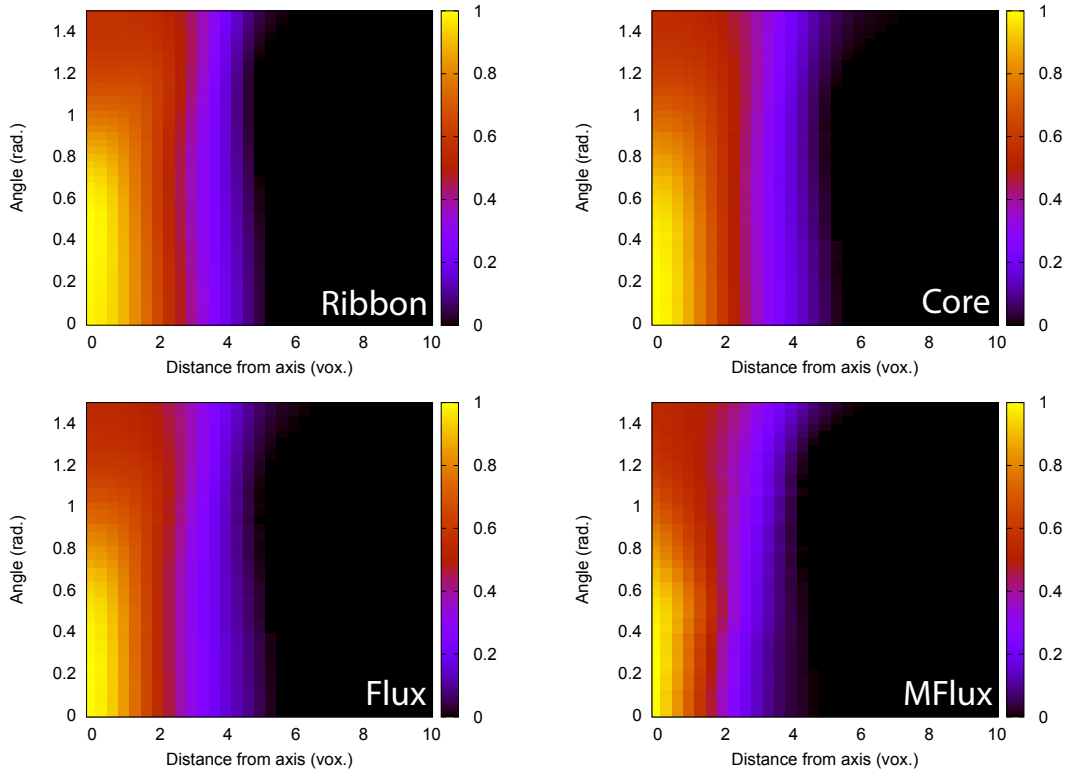


Figure 5.15: Joint variations of orientation and position. Test for a bar-convolved straight cylinder of radius 4. Scale is fixed to its optimal value. Top left: `Ribbon` . Top right: `Core` . Bottom left: `Flux` . Bottom right: `MFlux` .

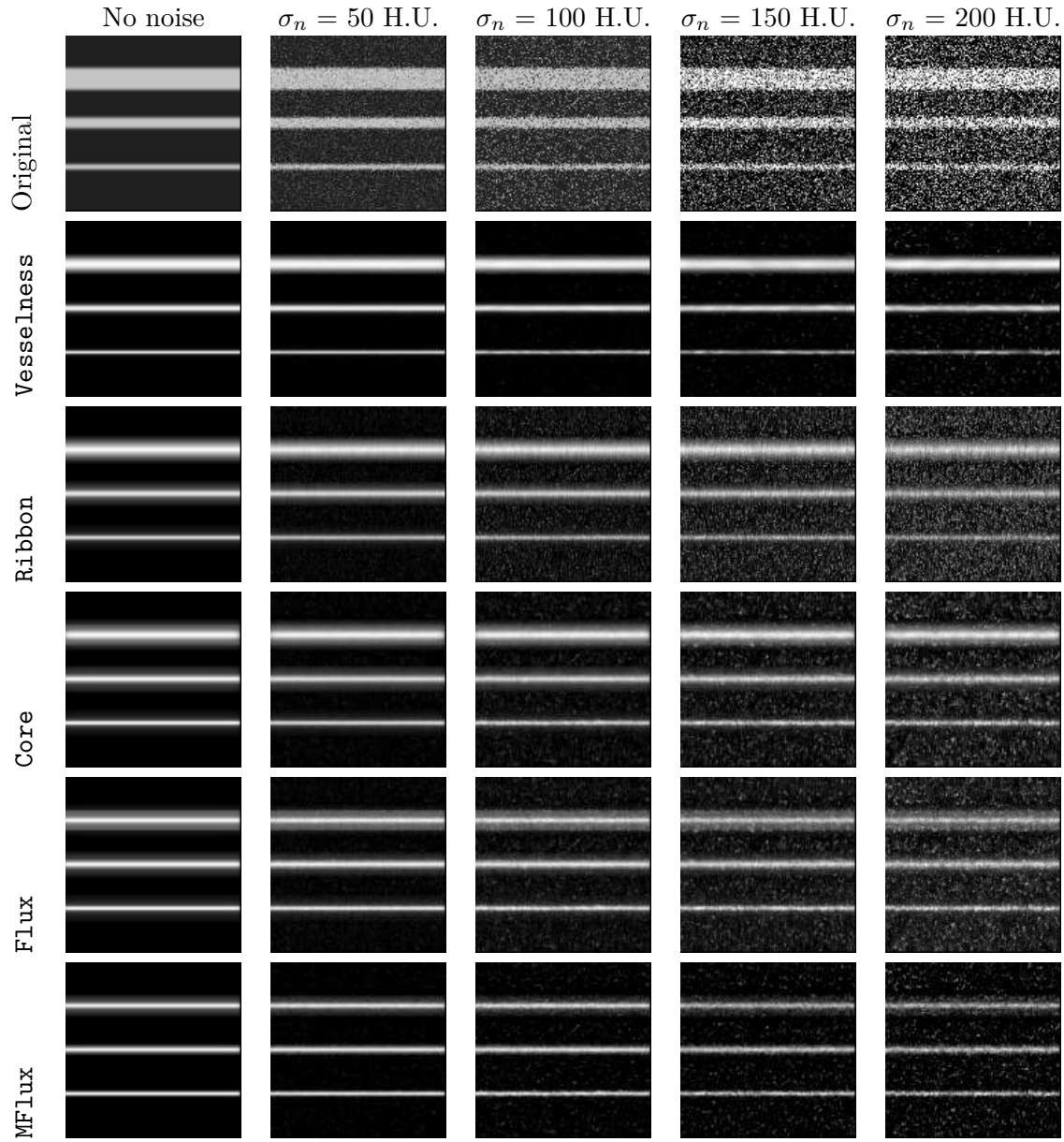


Figure 5.16: Sample responses in function of noise (Gaussian noise of standard deviation σ_n). Pixel intensities correspond to the maximum response over scales, for a single test orientation $(1, 0, 0)$.

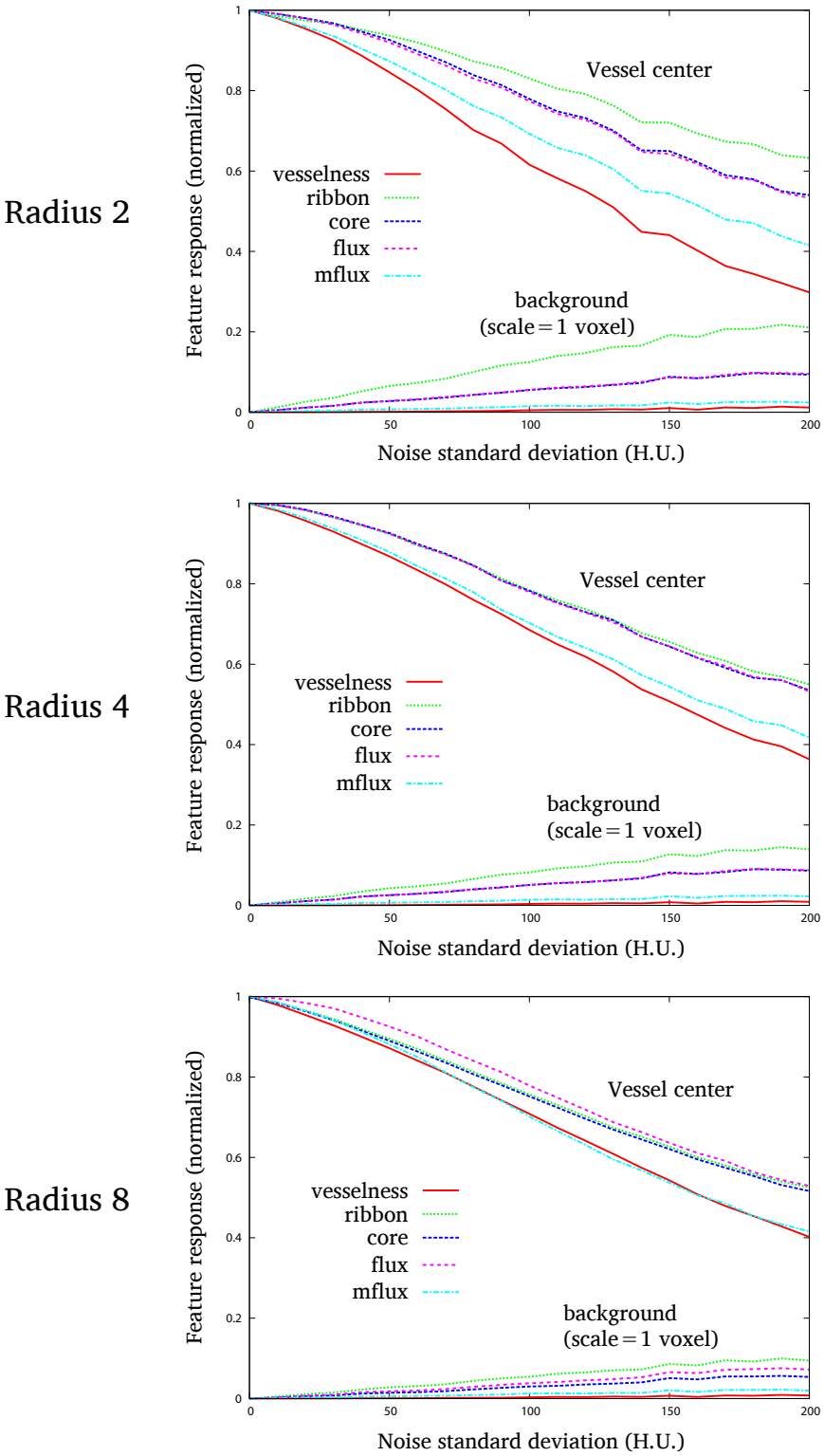


Figure 5.17: Average responses in function of noise (Gaussian noise of varying standard deviation σ_n) for the optimal parameters (center location, optimal radius and orientation) and in the background (radius 1 voxel, random orientation). From top to bottom: straight cylinders (contrast of 300 H.U.) of radius 2, 4 and 8 voxels, respectively. Results averaged for 1000 noise realizations.

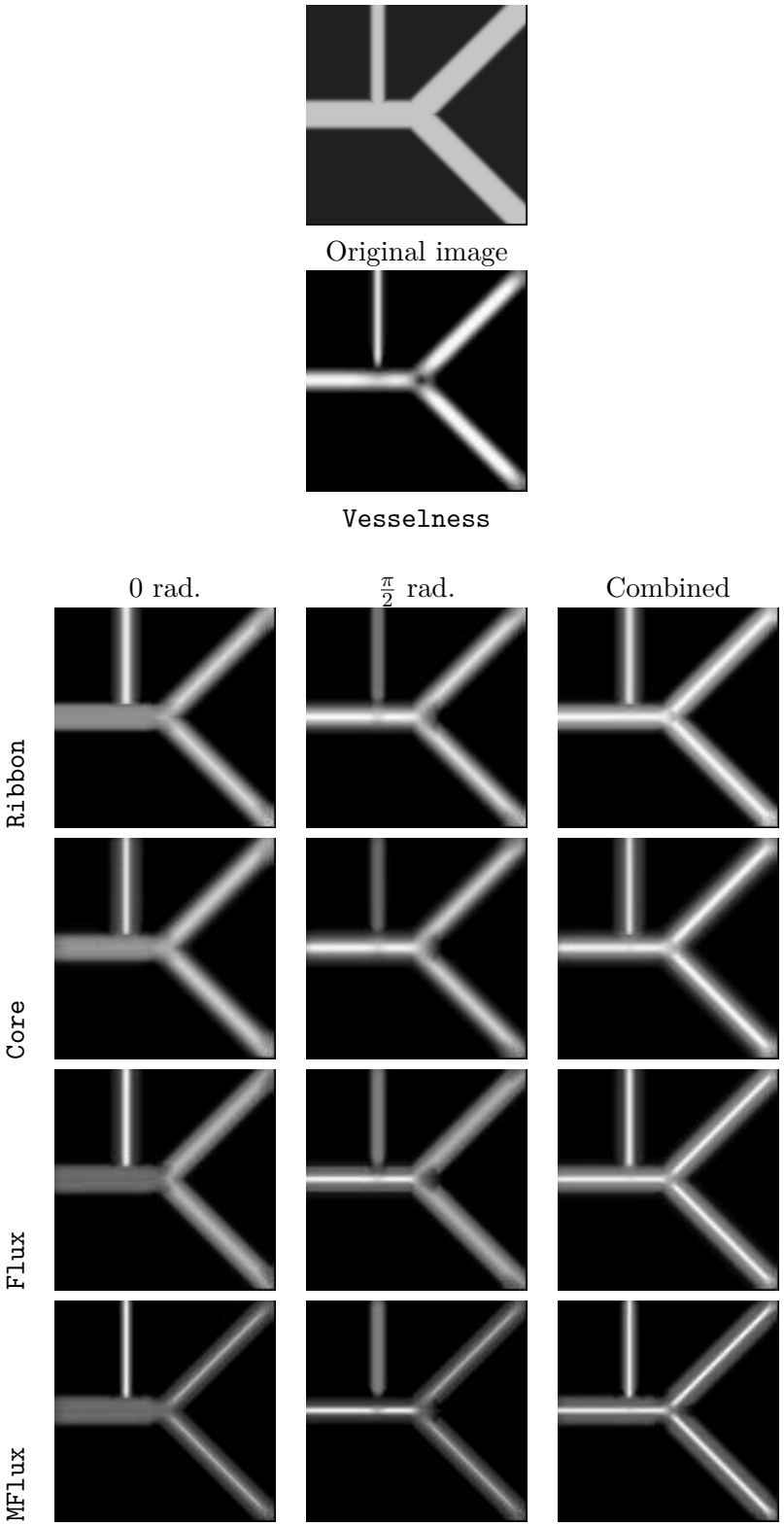
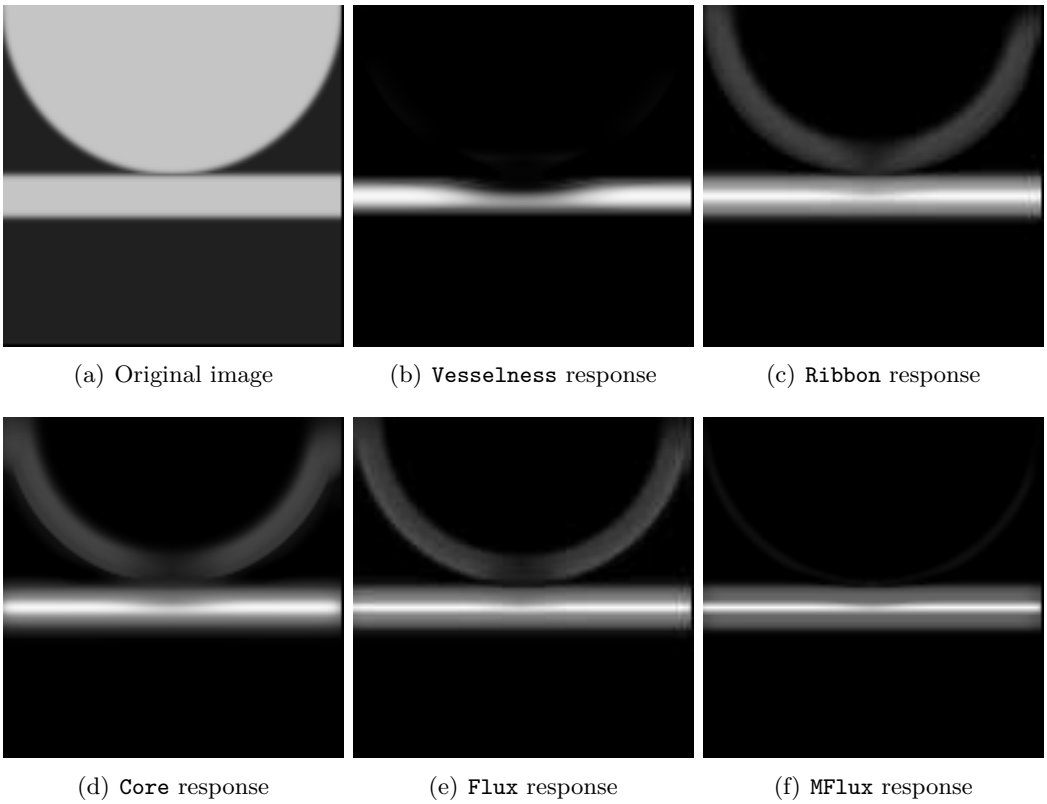


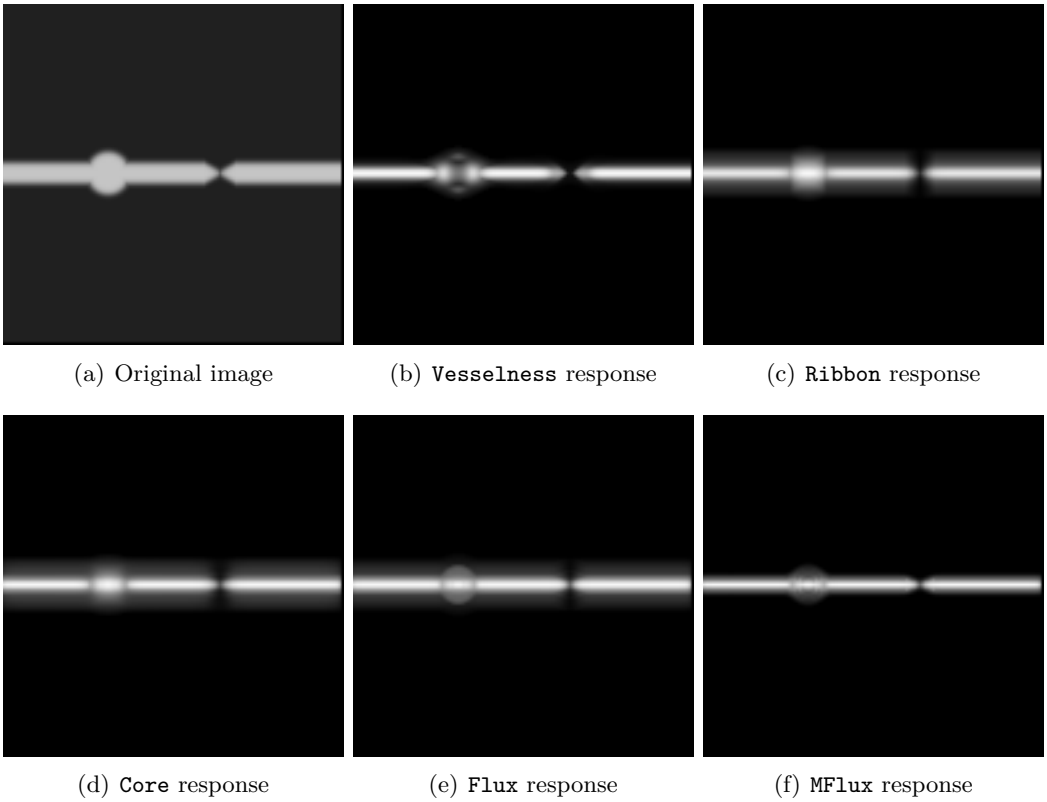
Figure 5.18: Branchings and orientation. Responses for an artificial vascular network consisting of a Y junction between vessels of radius 8 voxels and a T junction with a smaller vessel of radius 4 voxels. *Combined* images correspond to the maximal response over radiuses and orientations. For `Ribbon` , `Core` , `Flux` and `MFlux` , oriented responses (maximum over scales) are given for the vertical (0 rad.) and horizontal ($\frac{\pi}{2}$ rad.) orientations.



(a) Original image (b) `Vesselness` response (c) `Ribbon` response

(d) `Core` response (e) `Flux` response (f) `MFlux` response

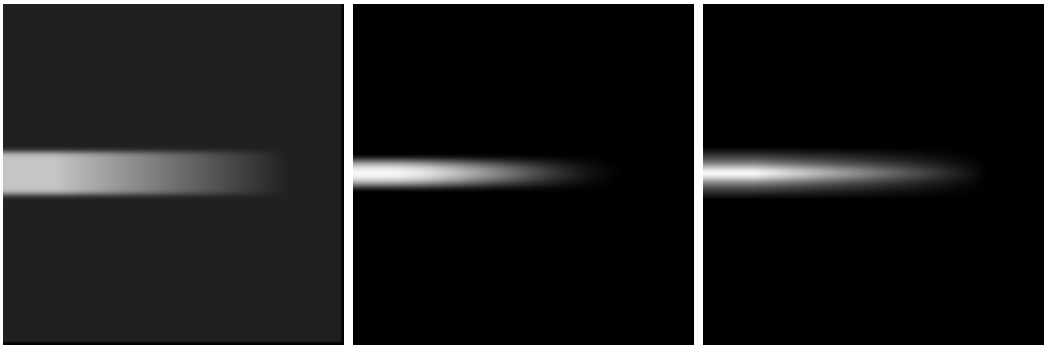
Figure 5.19: Influence of nearby structures and step edge response. Responses for a vessel of radius 8 voxels running nearby ($\Delta = 2$ voxels) a large-scale hyper-intense structure (orthogonal tube of radius 64 voxels). Pixel intensities correspond to the maximal response over radiuses and orientations.



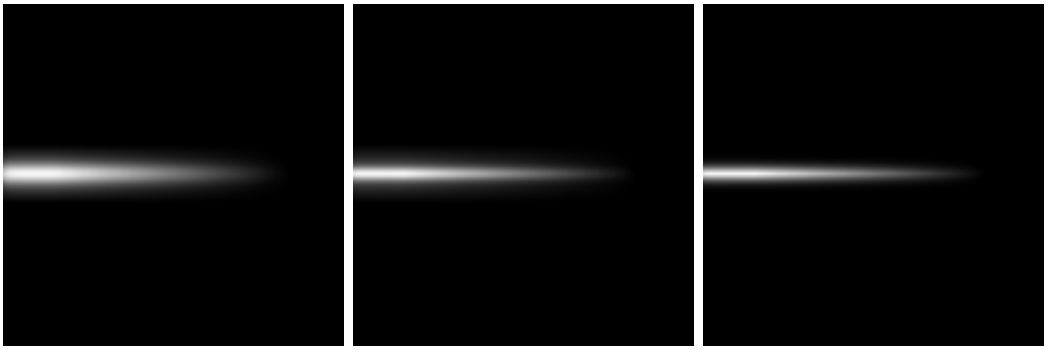
(a) Original image (b) `Vesselness` response (c) `Ribbon` response

(d) `Core` response (e) `Flux` response (f) `MFlux` response

Figure 5.20: Behavior with regards to local anomalies. Responses for a straight vessel ($r = 4$ voxels) with an artificial aneurysm (sphere of radius 8 voxels) and an artificial stenosis (total occlusion). Responses are given, per voxel, as the maximum response over radiuses, for a single test orientation $(1, 0, 0)$.



(a) Original image (b) `Vesselness` response (c) `Ribbon` response



(d) `Core` response (e) `Flux` response (f) `MFlux` response

Figure 5.21: Illustration of responses in function of the contrast. Responses given for a single vessel of radius 8 voxels, with contrast varying linearly from 300 to 0 H.U. Responses for a single radius (8 voxels) and single orientation $(1, 0, 0)$

.

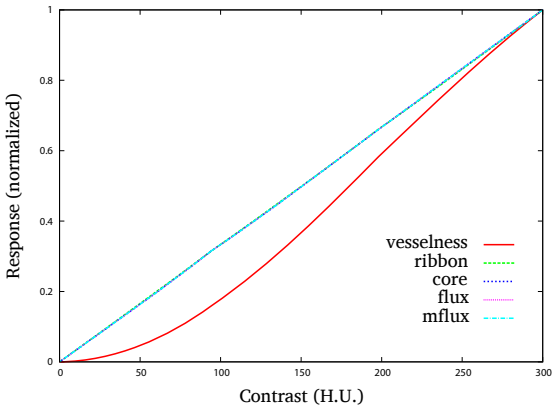


Figure 5.22: Responses in function of the contrast. Responses for optimal parameters (position, scale, direction), with contrast varying linearly from 0 to 300 H.U.

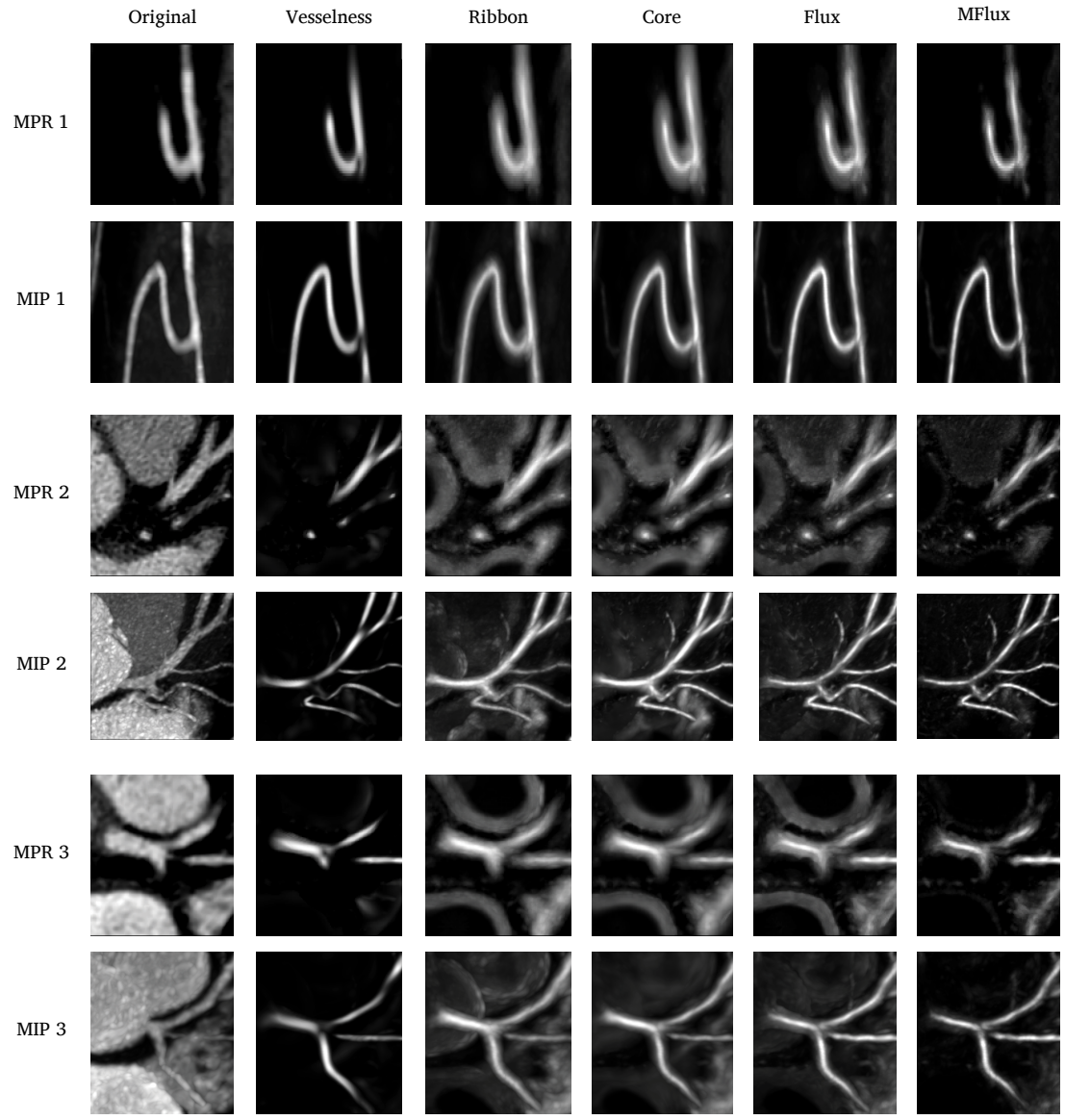


Figure 5.23: Details of feature responses on real cardiac CTA data. Pixel intensities correspond to the maximal response over radiuses and orientations. Results are given for 3 different datasets, with Multi-Planar Reformation 2D images (MPR) at the center of the volume of interest, and Maximum Intensity Projection (MIP) images.

Figure 5.24: `MFlux` response on whole cardiac CTA data: anterior 3D views. `MFlux` responses are given below each 3D rendering of the original data. Intensities correspond to the maximal response over radiuses and orientations.

Figure 5.25: MFlux response on whole cardiac CTA data: top 3D views. MFlux responses are given below each 3D rendering of the original data. Intensities correspond to the maximal response over radiuses and orientations.

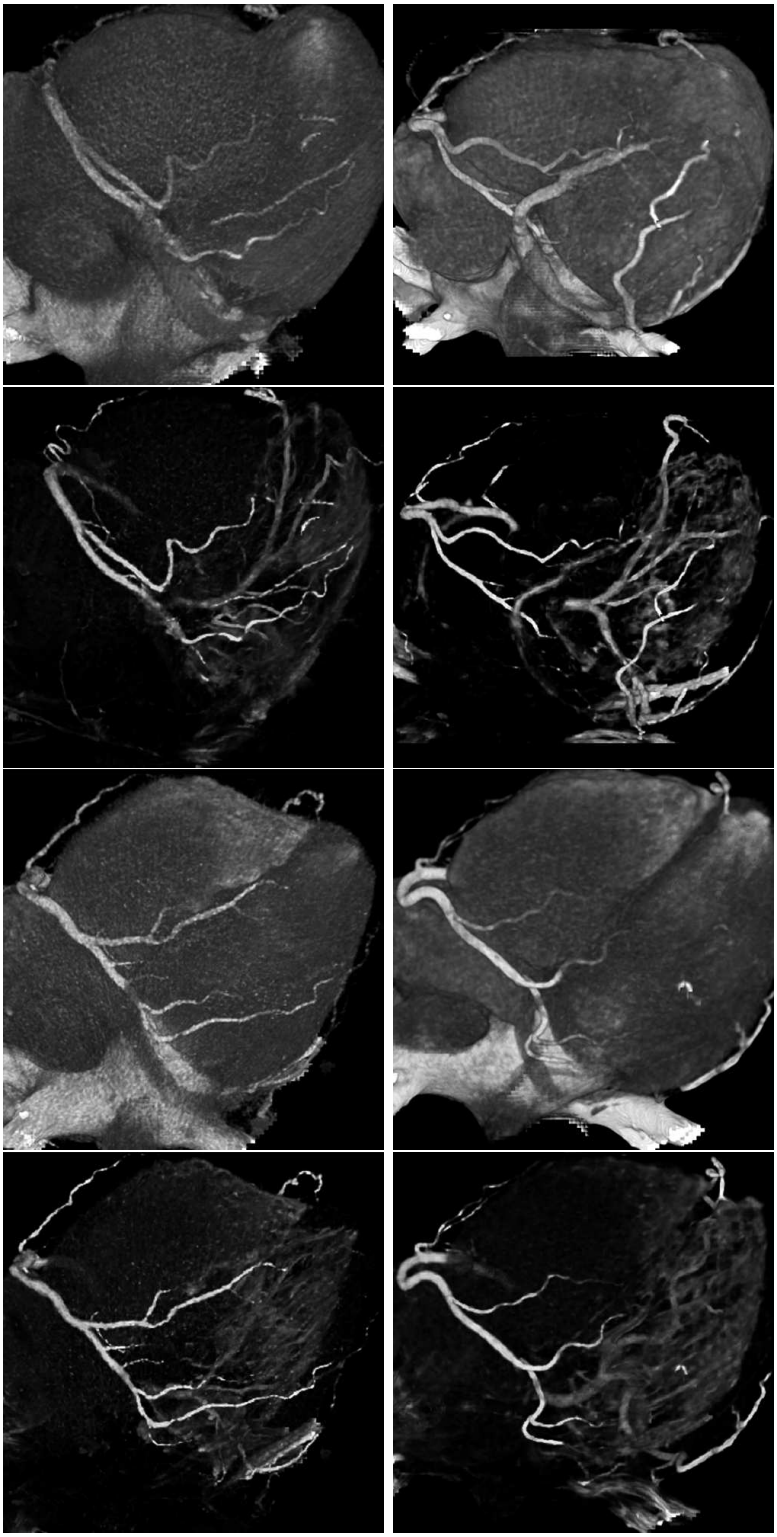


Figure 5.26: `MFlux` response on whole cardiac CTA data: bottom 3D views. `MFlux` responses are given below each 3D rendering of the original data. Intensities correspond to the maximal response over radiuses and orientations.

Chapter 6

Bayesian Model for 3D CT Coronaries

In this chapter, we detail our Bayesian model, which embeds prior knowledge on the variations of the geometric model (Chapter 4) and interprets feature responses (Chapter 5) in terms of likelihood distributions. We discuss the theoretical and practical definition of the different terms of the model, likelihood and priors.

6.1 Recursive Bayesian Formulation

Our geometric model (Chapter 4) defines vessels as discrete chains of states $x_t = (p_t, r_t, d_t)$ with p_t centerline points, r_t radius values and d_t local tangential directions. We now consider the model $x_{0:L}$ as a random variable. A particular chain $x^i_{0:L} = \{(p^i_t, r^i_t, d^i_t)\}^L_{t=0}$ then corresponds to a *realization* of this random variable.

A classical way of formulating a segmentation or detection task is through the maximization of the posterior probability of the model. The so-called Maximum A Posteriori (MAP) problem corresponds to the extraction of the realization of highest probability given the data:

$$x^*_{0:L} = \arg\max_{\mathcal{X}_L} p(x_{0:L}|z_{1:L}) \tag{6.1}$$

where $z_{1:L} = \{z_t\}^L_{t=1}$ are observation data available at each step t, derived from the image as discussed in Sec. 6.3. By convention, we assume that the initial state x_0 is given and fixed, at least partially, and that the first data measurements are obtained for $t = 1$. In our application, initialization is typically obtained from a user-provided or automatically detected position p_0. Unknown parameters (direction and radius) are uniformly distributed among a set of initial realizations $\{x^i_0\}$.

We emphasize that our applicative goals diverge from the classical MAP problem on different levels:

- the length of the target vessel is unknown;
- there might be several vessels of interest (branching structures).

A first remark is that in the general case, the search space explodes exponentially with the length of the chain. In practice, one can only explore a very limited fraction of this search space. Also, as we ultimately aim at extracting vascular networks comprising branchings, we are potentially interested in more than one result realization. Our main focus is thus on using the posterior probability to compare and select model realizations.

The posterior distribution is given by Bayes' theorem:

$$p(x_{0:L}|z_{1:L}) = \frac{p(x_{0:L})p(z_{1:L}|x_{0:L})}{p(z_{1:L})} \tag{6.2}$$

where $p(x_{0:L})$ is the *prior* distribution and $p(z_{1:L}|x_{0:L})$ the joint *likelihood* of the observations. The prior probability is a model-based term, which integrates prior knowledge on the variations of the model. The likelihood term can be thought of as a data attachment term. The *evidence* term $p(z_{1:L})$ is independent of the model realization, but is linked to the length of the model.

In the extraction strategies discussed in Chapters 7 and 8, the optimization of the model is carried out in an iterative fashion through tracking/propagation approaches. It is thus interesting to express the posterior distribution in a recursive form. First, assuming a first order Markovian transition model, the joint prior becomes:

$$p(x_{0:L}) = p(x_0)\prod_{t=1}^{L} p(x_t|x_{t-1}) \tag{6.3}$$

Second, we assume that the observation sets $\{z_t\}$ are conditionally independent given the model:

$$p(z_{1:L}|x_{0:L}) = \prod_{t=1}^{L} p(z_t|x_t) \tag{6.4}$$

These assumptions are classically adopted for sequential Bayesian estimation problems (Doucet et al., 2000). In particular, they allow the expression of the posterior in a recursive form.

Combining Eq. 6.2, 6.4 and 6.3, one can derive a recursive form of the posterior (Doucet et al., 2000):

$$\begin{aligned}
p(x_{0:t}|z_{1:t}) &= \frac{p(x_{0:t})p(z_{1:t}|x_{0:t})}{p(z_{1:t})} \\
&= p(x_0)\prod_{k=1}^{t}\frac{p(x_k|x_{k-1})p(z_k|x_k)}{p(z_{1:t})} \\
&= \frac{p(x_t|x_{t-1})p(z_t|x_t)}{p(z_t|z_{1:t-1})}p(x_0)\prod_{k=1}^{t-1}\frac{p(x_k|x_{k-1})p(z_k|x_k)}{p(z_{1:t-1})} \\
&= \frac{p(x_t|x_{t-1})p(z_t|x_t)}{p(z_t|z_{1:t-1})}p(x_{0:t-1}|z_{1:t-1})
\end{aligned} \tag{6.5}$$

The posterior is updated from time $t-1$ to t through the transition prior $p(x_t|x_{t-1})$ and marginal likelihood $p(z_t|x_t)$, which suffice to define our Markovian process. The remaining of the chapter is dedicated to the definition of these terms for our CTA coronary application. Last but not least, the term $p(z_t|z_{1:t-1})$, relative to the data innovation, corresponds to the adjustment of the normalization with regards to the length of the model. It is generally impossible to solve this term analytically. A classical exception is the case of linear Gaussian state space models, for which Kalman filtering (Kalman, 1960) provides an analytical solution of the innovation process. As will be illustrated in the following sections, our model is inherently non-Gaussian and non-linear. In Chapter 7, the segmentation task is reformulated as a minimal cumulative path problem, where the normalization issue is handled heuristically. In Chapter 8, numerical integration is made possible by the use of point-wise Monte-Carlo estimations, in a so-called particle filtering framework.

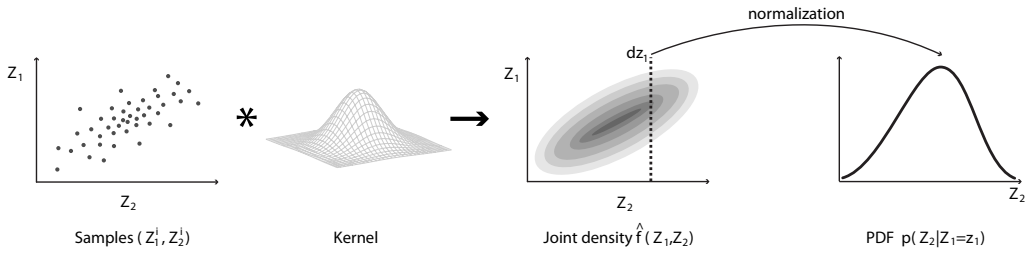


Figure 6.1: Overview of the density estimation process.

6.2 Learning from the Ground-Truth Database

As discussed in Appendix B, we have access to a ground-truth database of 61 manually delineated cardiac CTA datasets, totaling 858 vessel segments with centerline and radius segmentation from experts. The models fitted to the ground-truth delineations (see Sec. 4.1.3) provide us with samples of various statistical variables, including successive scales, directions, and `MFlux` responses associated to true vessel parameters. We exploit this valuable source of statistical information by learning the distributions composing Eq. 6.5 thanks to a non-parametric density estimation scheme.

Implicitly, we assume that the probability *distributions* composing our model admit probability *density* functions (PDFs). In the remaining of the document, we abusively assimilate the distributions and their PDFs. In effect, we propose to learn the PDFs associated to these distributions.

Considering two random variables Z_1 and Z_2, our general estimation scheme can be described as follows:

1. estimate the joint density $\hat{f}(Z_1, Z_2)$ from samples $\{(Z_1^i, Z_2^i)\}$;
2. extract conditional probability densities of the form $p(Z_2|Z_1 = z_1)$.

An overview of this scheme is given in Fig. 6.1. Intuitively, densities $\hat{f}(Z_1, Z_2)$ capture the joint variations of Z_1 and Z_2, for instance the joint variations of successive radiuses, for $Z_1 = r_{t-1}$ and $Z_2 = r_t$. We propose to learn these joint densities from the samples at our disposal, using non-parametric kernel estimation (Silverman, 1986). Given these joint densities, we can eventually extract any conditional PDF needed by our model. This general principle is applied to both the likelihood and prior parts of our model, whose respective decompositions are discussed in Sec. 6.3 and Sec. 6.4.

6.2.1 Learning Joint Densities by Kernel Density Estimation

The general formulation of kernel density estimation is as follows. Given n data points x^i (samples of a d-dimensional space R^d), the density at point x is estimated as:

$$\hat{f}(x) = \frac{1}{n} \sum_{i=1}^{n} K_H(x - x^i)$$

with

$$K_H(x) = |H|^{-1/2} K(H^{-1/2} x)$$

$K(x)$ is a $d-$variate kernel and H designates a symmetric positive definite $d \times d$ *bandwidth* matrix. In the case where the underlying density is known to be Gaussian, optimal kernel

and bandwidth choices (with regards to the mean integrated squared error, MISE criterion) can be analytically derived (Silverman, 1986). In the general, multivariate case, these choices are more intricate (Silverman, 1986; Wand and Jones, 1995). Various schemes to automatically select the bandwidth parameters have been proposed, depending on the final application (estimation, clustering, filtering) (Comaniciu, 2003).

For the sake of simplicity, we empirically chose a two-dimensional, uncorrelated Gaussian kernel (H diagonal). The estimated joint density $\hat{f}(Z_1, Z_2)$ of two random variables Z_1 and Z_2 is then given by:

$$\hat{f}(Z_1, Z_2) = \frac{1}{nh_1h_2} \sum_{i=1}^{n} K_g(\frac{Z_1 - Z_1^i}{h_1}, \frac{Z_2 - Z_2^i}{h_2})$$

with

$$K_g(x, y) = \frac{1}{2\pi} exp(-\frac{1}{2}(x^2 + y^2))$$

where (Z_1^i, Z_2^i) are n samples from the joint density, and h_1 and h_2 are bandwidth parameters associated to Z_1 and Z_2, respectively, and n the number of samples. We experimented with different kernels, such as the Epanechnikov kernel, without significant differences in behavior.

Thanks to this scheme, we are able to estimate rich, continuous joint densities from our database of discrete samples. This estimation is said to be non-parametric as no distribution model is imposed. In the following sections, bandwidth parameters h_1 and h_2 will be specified empirically, depending on the variables Z_1 and Z_2 considered. They are always chosen so as not to over-smooth the resulting densities and preserve the subtle variations embedded in the training database.

6.2.2 Extraction of Conditional Probability Density Functions

Considering the joint density $\hat{f}(Z_1, Z_2)$ as a joint probability density $p(Z_1, Z_2)$ (after normalization), we can extract any conditional PDF of the form $p(Z_2|Z_1 = z_1)$ following Bayes' rule:

$$p(Z_2|Z_1 = z_1) = \frac{p(Z_2, Z_1 \in dz_1)}{p(Z_1 \in dz_1)} = \frac{p(Z_2, Z_1 \in dz_1)}{\int p(Z_2 \in dz_2, Z_1 \in dz_1) dz_2} \tag{6.6}$$

The notation $p(Z_2|Z_1 = z_1)$ indicates that conditioning is performed with regards to a certain realization z_1 of Z_1. Notations dz_1 and dz_2 correspond to infinitesimal volumes around realizations z_1 and z_2 on which continuous distributions are locally integrated.

In practice, we use such conditional PDFs to evaluate directly the different terms of our model.

6.3 Marginal Likelihood $p(z_t|x_t)$

This section details the decomposition of observation model for CTA coronaries, *i.e.*, the definition of the marginal likelihood $p(z_t|x_t)$ used in our recursive model (Eq. 6.5). At each step t, $z_t = \{y_t^j\}_{j=1}^{N_t}$ corresponds to a *set* of available observations. An observation y_t^i is obtained as the response of the `MFlux` feature (Chapter 5) linked to a state realization x_t^i. In other words, an observation y_t^i is obtained as the `MFlux` response for position p_t^i, radius r_t^i and direction d_t^i. Notice that in our context, the image is static, but the number of

possible observations is tremendous (actually infinite). A particularity of our applicative framework is that we control the source of the observations (the image) and can actually generate as many observations as we desire. For the sake of efficiency, we will naturally generate observations only for the model realizations $\{x^i_{1:L}\}$ considered by the extraction schemes (Chapters 7 and 8), so that an observation is always linked to a state realization x^i_t.

The term $p(z_t|x_t)$ thus corresponds to the joint likelihood of all available observations, given the model state at time t. First, we assume that observations y^j_t are conditionally independent given the model:

$$p(z_t|x_t) = \prod_{j=1}^{N_t} p(y^j_t|x_t) \tag{6.7}$$

This assumption is admittedly inexact, as the supports of different `MFlux` evaluations may locally overlap. This approximation has the key advantage of significantly simplifying further developments, providing us with a tractable model. Following principles previously proposed in (Geman and Jedynak, 1996; Konishi et al., 2003; Yuille and Coughlan, 1997, 2000), a particularity of our observation model is to consider distinct *vessel* and *background* likelihood distributions. Observations can either be generated by actual vessels with distribution $p_v(y_t|r_t = r^i_t)$, or by the background with distribution $p_{bg}(y_t)$. As will be illustrated in Chapters 7 and 8, explicitly taking information about the background has advantageous consequences in terms of selectivity and robustness. We exploit the fact that observations are linked to model states when evaluating the likelihood value given a model state realization x^i_t:

$$p(y^j_t|x_t = x^i_t) = \begin{cases} p_v(y^i_t|x_t = x^i_t) & \text{if } j = i \\ p_{bg}(y^j_t) & \text{otherwise} \end{cases} \tag{6.8}$$

Given a state realization x^i_t, an observation y^j_t is considered as coming from a vessel distribution with probability $p_v(y^i_t|x_t = x^i_t)$ if it is linked to the state x^i_t ($j = i$). If y^j_t is not linked to x^i_t ($j \neq i$), it is valued as coming from the background distribution with probability $p_{bg}(y^j_t)$. We emphasize that vessel likelihood PDFs $p_v(y_t|x_t)$ are dependent on the model. More precisely, they vary with the radius r_t considered, as will be detailed thereafter. On the other hand, the background distribution $p_{bg}(y_t)$ is independent of the model.

One can now decompose the marginal likelihood value given a model realization as:

$$p(z_t|x_t = x^i_t) = p_v(y^i_t|x_t = x^i_t) \prod_{j \neq i} p_{bg}(y^j_t) \tag{6.9}$$

From Eq. 6.9, one can see that the valuation of the likelihood is dependent on the model state realizations. In other words, different model realizations lead to different interpretations of the set of observations. In that sense, realizations compete for the best explanation of all the observations. One can re-factor Eq. 6.9 as follows:

$$p(z_t|x_t = x^i_t) = \frac{p_v(y^i_t|x_t = x^i_t)}{p_{bg}(y^i_t)} \prod_{y^j_t \in z_t, \forall j} p_{bg}(y^j_t) \tag{6.10}$$

The shift from Eq. 6.9 to Eq. 6.10 has important implications. The first term in Eq. 6.10 is a likelihood ratio between vessel and background for the single observation

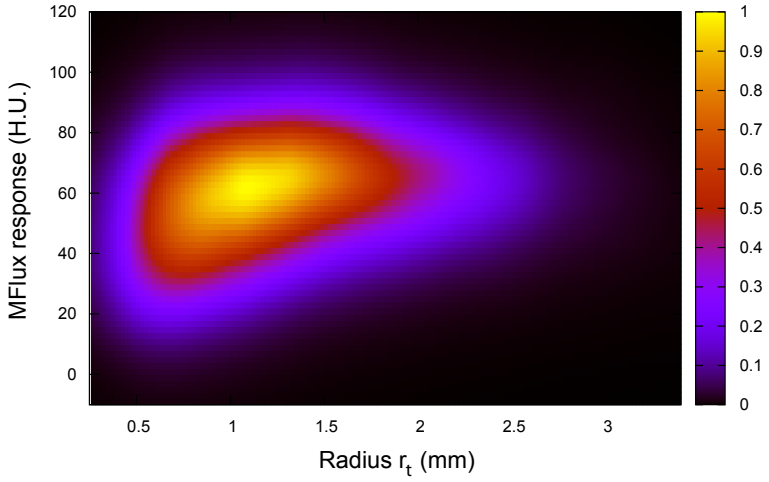


Figure 6.2: Unnormalized joint density $\hat{f}_v(y_t, r_t)$ of observations y_t (`MFlux` responses) for true vessels over radiuses r_t. The density is obtained from samples (y_t^i, r_t^i), using kernel estimation with bandwidth parameters $h_{y_t} = 10$ H.U. and $h_{r_t} = 0.3mm$, chosen so as to obtain a smooth density landscape without over-smoothing. Result learned from 10 datasets randomly selected in our database.

y_t^i linked to the state realization x_t^i. This ratio directly encodes the competition between vessel and background distributions for the interpretation of the observation. The second term is a product over all the observations composing $z_t = \{y_t^j\}_{j=1}^{N_t}$ and is independent of the model. This allows us to focus solely on the observations associated to the model realization at hand.

6.3.1 Vessel Likelihood $p_v(y_t|x_t)$

The geometric models fitted to the ground-truth segmentations give us samples of true vessel parameters (position, radius, direction), for which we can evaluate corresponding `MFlux` feature responses. We assume the invariance by rotation and translation of observations y_t, that is $p_v(y_t, x_t) = p_v(y_t, r_t)$. Using the principles presented in Sec. 6.2, we estimate the joint density $\hat{f}_v(y_t, r_t)$, that is, the joint density of `MFlux` responses for true vessels over radiuses (Fig. 6.2).

From the joint density, we are able to extract likelihood PDFs of the form $p_v(y_t|x_t = x_t^i) = p_v(y_t|r_t = r_t^i)$ by considering a particular radius realization r_t^i. A selection of such density functions is given in Fig. 6.4. Our non-parametric density estimation is able to capture subtle variations of `MFlux` responses as a function of the target radius. In particular, one can note in Fig. 6.4 that responses tend to be lower, and of larger variance, for smaller vessels. This observation can be explained by the loss of contrast affecting small-scale coronaries.

6.3.2 Background Likelihood $p_{bg}(y_t)$

To construct the background likelihood, one could use a strategy similar to that employed for the vessel likelihood in the previous section, specifically evaluating `MFlux` responses at parameters (locations, scales, directions) *not* corresponding to delineated vessels

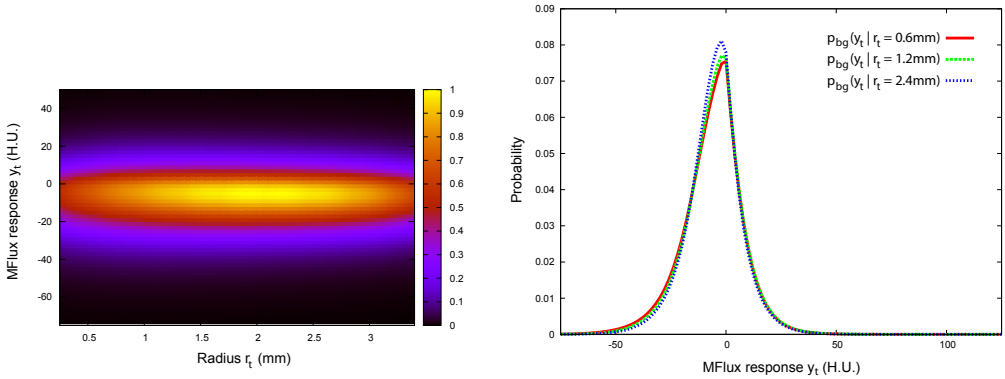


Figure 6.3: Background likelihood: independence with regards to the model. Left: for a selected dataset, joint density $\hat{f}_{bg}(y_t, r_t)$ estimated in the background ($h_{y_t} = 10$ H.U. and $h_{r_t} = 0.3mm$, see Sec. 6.2). Right: background likelihood PDFs $p_{bg}(y_t|r_t = r_t^i)$ for different radiuses r_t^i.

in our ground-truth database. We use a slightly different approach, not relying on any ground-truth knowledge, as proposed in (Konishi et al., 2003). We exploit the fact that on any given cardiac CTA dataset, randomly selected parameters are very unlikely to correspond to coronaries. The reader may refer to (Konishi et al., 2003) for a more formal analysis of the error induced by such a scheme, in the context of edge detection. Of particular interest, it is demonstrated in (Konishi et al., 2003) that the information lost by this approximation is negligible.

In Sec. 6.3, we mentioned that we assume that the background likelihood $p_{bg}(y_t)$ is independent of the model realization. This assumption, which is admittedly an approximation, relies on the analysis of background responses on different datasets. Besides rotation and translation invariance, we evaluated the behavior of background responses with regards to the tested radius. In Fig. 6.3, we illustrate the joint density $\hat{f}_{bg}(y_t, r_t)$ and different likelihood PDFs $p_{bg}(y_t|r_t = r_t^i)$ estimated for a selected dataset. It can be seen that background likelihood density functions are almost independent of the tested radius. This observation holds for all the datasets in our database.

On the other hand, we capture the variations in background likelihood PDFs between datasets. As can be seen in Fig. 6.4, background likelihood distributions vary mainly with the noise level of the acquisition at hand. In practice, our final estimation scheme of the data-dependent background likelihood is straightforward. Given a dataset, we randomly sample 10^5 parameters (p_t^i, r_t^i, d_t^i) for which we evaluate the response of the `MFlux` feature. This number proved to be high enough to provide us with sufficient, stable statistics, and requires less than 0.1 sec. to be evaluated on a 2.1GHz processor (see Sec. 5.3). Our sampling scheme excludes hypo-intense areas such as lung areas. The histogram of the responses obtained is then smoothed with a Gaussian of standard deviation $\sigma_{y_t} = 10$H.U. and normalized to obtain the likelihood PDF $p_{bg}(y_t)$.

6.3.3 Semantics, Overlap and Ambiguity

As our goal is to maximize Eq. 6.5, the likelihood values can be interpreted as score mappings, turning `MFlux` feature responses into multiplicative, probabilistic scores. As

expected, vessel likelihood PDFs $p_v(y_t|r_t)$ concentrate their mass for relatively high responses. Conversely, background likelihoods $p_{bg}(y_t)$ concentrate for relatively low responses. Following a frequency-based point of view, these likelihood densities penalize rare observations, with regards to the initial set of samples from their respective classes. In other words, the observations of highest likelihood values are the ones most frequently recorded for the vessel and background classes.

A first important remark is that our likelihood model implicitly exploits the fact that our CTA data are *calibrated*. We make the assumption that CTA coronaries have relatively stable contrast properties. Responses of our `MFlux` feature are highly dependent on the local contrast, and the vessel likelihood distributions (Fig. 6.4) indeed exhibit single concentrated modes. Such assumptions would not hold for other, non-calibrated angiographic imaging modalities such as MRA.

Second, it is worth noting that the maximization of the likelihood is not equivalent to a direct maximization of the feature response, as responses of maximum vessel and background likelihoods do not correspond to highest and lowest feature responses, respectively. For instance, a very high `MFlux` response can potentially become "too good to be true" and be penalized in terms of vessel likelihood. Conversely, a very low `MFlux` response can result in a relatively low background likelihood value. This can theoretically lead to counter-intuitive effects, where realizations with lower (but still relatively high) responses are preferred to "too good" ones. In our experiments, we did not witness obvious manifestations of these potential issues. This seems to confirm that realizations of highest likelihood coincide relatively closely with local maxima of the feature response. It may be worth studying more precisely the actual impact of this likelihood formulation on the accuracy of the extracted result. In our opinion, these potential drawbacks are outweighed by the advantages offered by our approach. Numerous vascular tracking scheme rely on an *ad-hoc* formulation of their cost/energetic functional from the feature response, effectively enforcing a monotonic behavior. In contrast, our approach loses the monotonic property, but the principal components of the posterior PDF are learned automatically from a training database. Our likelihood formulation allows for the integration of background information and is adaptive with regards to the target radius (because of $p_v(y_t|r_t)$) and the data at hand (because of $p_{bg}(y_t)$). We believe that such properties contribute largely to the robustness of the extraction we propose in Chapters 7 and 8.

From Eq. 6.10, one can note that vessel and background likelihood values are used in a ratio form, *i.e.* are compared to each other. As discussed in (Yuille and Coughlan, 2000) and (Konishi et al., 2003), the complexity of such a Bayesian inference problem is driven by different *order* parameters, one of which being formulated on the object and background likelihood densities. More precisely, the important factor is the separability of these densities, encoding the intrinsic ambiguity of the data. The reader may refer to (Yuille and Coughlan, 2000) and (Konishi et al., 2003) for discussions on how this ambiguity, expressed in terms of detection error rates, can be estimated through Bhattacharrya coefficients and Chernoff information criteria. Loosely speaking, these criteria are functions of the overlap between the background and object likelihoods. As can be seen in the examples of Fig. 6.4, the likelihood densities for our coronary application are generally well separated, confirming the good discriminative power of the `MFlux` feature. Their overlap vary, however, as function of the target radius (because of $p_v(y_t|r_t = r_t^i)$) and the data at hand (because of $p_{bg}(y_t)$). Finally, we emphasize that our segmentation problem does not reduce to a decision issue based on likelihood ratios. Likelihood information is coupled with *prior* terms to guide and constrain the Bayesian model. The definitions of

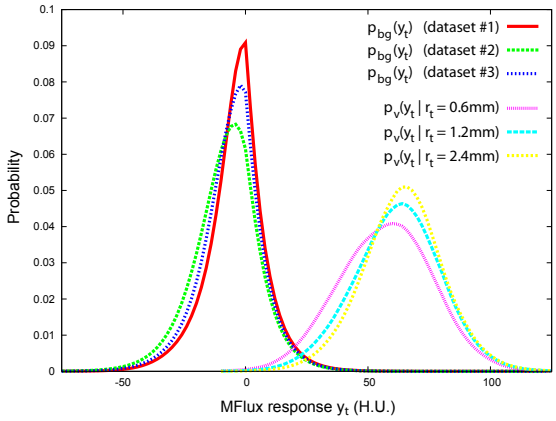


Figure 6.4: Samples of vessel and background likelihood PDFs. Vessel likelihood PDFs $p_v(y_t|r_t = r_t^i)$, for different radius values r_t^i, are extracted from the joint density depicted in Fig. 6.2. Background likelihoods $p_{bg}(y_t)$ are independent of the model, but are specific to a given dataset (see Sec. 6.3.2).

these prior terms are discussed in the following sections.

6.4 Transition Prior $p(x_t|x_{t-1})$

The last term from Eq. 6.5 requiring definition is the first order Markovian transition *prior* $p(x_t|x_{t-1})$. It corresponds to prior knowledge on the potential variations of the geometric model. We express the transition prior as a function of radius and direction variations:

$$p(x_t|x_{t-1}) = p(p_t, r_t, d_t|p_{t-1}, r_{t-1}, d_{t-1}) = p(r_t|r_{t-1})p(d_t|d_{t-1}, r_{t-1}) \tag{6.11}$$

To allow this decomposition, we assume a shift-invariant prior, that is, the prior does not depend on the centerline position p_t. We additionally assume that scale variations are independent of the direction. The transition prior is thus composed of two terms, the scale transition prior $p(r_t|r_{t-1})$ and the direction prior $p(d_t|d_{t-1}, r_{t-1})$.

6.4.1 Scale Transition Prior $p(r_t|r_{t-1})$

Scale transition PDFs $p(r_t|r_{t-1} = r_{t-1}^i)$, depending on the realization r_{t-1}^i (previous radius), are extracted from the joint density $\hat{f}(r_t, r_{t-1})$. This joint density is learned from successive radius samples (r_t^i, r_{t-1}^i) from the ground-truth-fitted models (see Sec. 6.2) and is illustrated in Fig. 6.5. It can be noted that areas of higher density lie along the axis $r_t = r_{t-1}$, confirming the intuition that scale variations along coronaries are relatively slow and that an artery retains locally a relatively stable radius. In Fig. 6.6, we give examples of transition priors $p(r_t|r_{t-1} = r_{t-1}^i)$ for selected radius realizations r_{t-1}^i. Slight differences can be seen, such as increased variance for larger vessels. In general, these distributions are not symmetric and not centered exactly around the previous scale r_{t-1}^i, their mode typically appearing for a slightly lower value. As is increasingly the case for larger vessels, it confirms the observation that coronaries tend to decrease in radius from the ostia to distal ends.

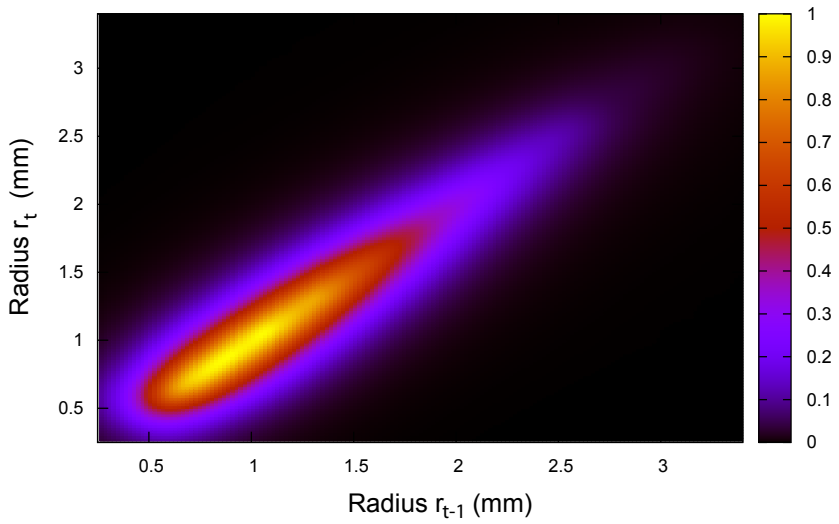


Figure 6.5: Unnormalized joint density $\hat{f}(r_t, r_{t-1})$ of successive radiuses. The density is obtained from samples (r_t^i, r_{t-1}^i), using kernel estimation with bandwidth parameters $h_{r_t} = h_{r_{t-1}} = 0.3mm$, chosen so as to obtain a smooth density landscape without over-smoothing. Result for a discretization step $s = 0.3mm$, learned from 10 datasets randomly selected in our database.

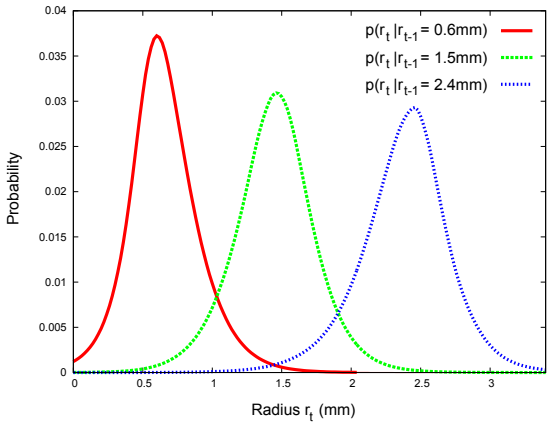


Figure 6.6: Samples of scale transition priors. PDFs of scale transition priors $p(r_t|r_{t-1} = r_{t-1}^i)$, for different radius values r_{t-1}^i, are extracted from the joint density depicted in Fig. 6.5 (see Sec. 6.3.1).

6.4.2 Direction Prior $p(d_t|d_{t-1}, r_{t-1})$

Intuitively, blood vessels have a propensity to follow relatively smooth trajectories. The directional term $p(d_t|d_{t-1}, r_{t-1})$ embeds such prior knowledge, defined through the tangential angular changes $\alpha_t = \arccos d_t \cdot d_{t-1} = \arccos \frac{(c_{t-1}-c_{t-2})\cdot(c_t-c_{t-1})}{\|c_{t-1}-c_{t-2}\|\|c_t-c_{t-1}\|}$ (Fig. 6.7):

$$p(d_t|d_{t-1}, r_{t-1}) = p(\alpha_t|r_{t-1}) \tag{6.12}$$

In spirit, this approach is similar to (Schaap et al., 2007b) and (Allen et al., 2008), who fixed parametric distributions, respectively Gaussian and von-Mises-Fisher. In contrast,

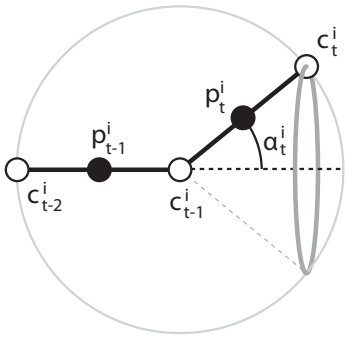


Figure 6.7: Direction prior and tangential angle. The angle $\alpha_t = \arccos d_t \cdot d_{t-1} = \arccos \frac{(c_{t-1}-c_{t-2})\cdot(c_t-c_{t-1})}{\|c_{t-1}-c_{t-2}\|\|c_t-c_{t-1}\|}$ corresponds to the angular deviation between two consecutive directions d_{t-1} and d_t.

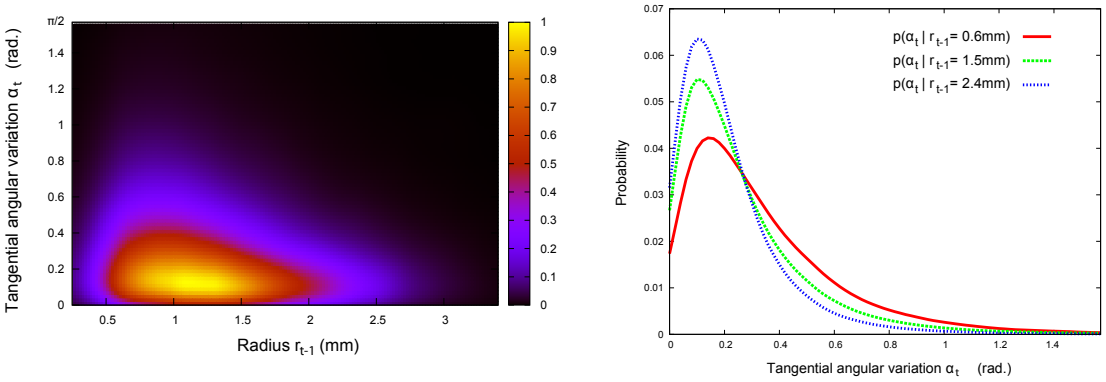


Figure 6.8: Estimation of the direction prior. Left: unnormalized joint density $\hat{f}(\alpha_t, r_{t-1})$ of angular variations and radiuses. The density is obtained from samples (α_t^i, r_{t-1}^i), using kernel estimation with bandwidth parameters $h_{\alpha_t} = 0.05rad$ and $h_{r_{t-1}} = 0.3mm$. Right: samples of PDFs $p(\alpha_t | r_{t-1} = r_{t-1}^i)$, for different radius values r_{t-1}^i, extracted from the joint density. Result for a discretization step $s = 0.3mm$, learned from 10 datasets randomly selected in our database.

we again learn the directional PDFs in a non-parametric fashion. Fig. 6.8 illustrates the joint density $\hat{f}(\alpha_t, r_{t-1})$ estimated from our database, along with examples of densities $p(\alpha_t | r_{t-1} = r_{t-1}^i)$ for different radius values. First, one can note that smaller vessels exhibit higher angular variations. Second, the maximal probability density is attained for a small, but non-zero angle. As natural structures, coronaries generally do not run exactly straight.

6.5 Discussion

In this chapter, we presented our recursive, Bayesian *a posteriori* model for CTA coronary arteries. One of the particularities of our model, inspired by the works of (Geman

and Jedynak, 1996; Yuille and Coughlan, 2000; Konishi et al., 2003), is the data likelihood formulation which integrates background information. We also discussed how we can learn, using non-parametric density estimation, the different PDFs of the distributions involved in our model. The principal contribution of this scheme is to enable the capture of subtle dependencies, such as the variations of the vessel likelihood in function of the target radius. We stress that this estimation process is general, flexible and suited for any effective discretization of the parameter space. We estimate *continuous* joint densities from which we can extract any conditional density needed.

A strong motivation behind our approach was the direct exploitation of the statistical information embedded in our ground-truth database, with a minimal recourse to empirical parameterization. A debatable counterpart is that the frequency-, sample-based approach underlying our learning procedure can have undesired counter-intuitive effects (*e.g.*, observations considered as "too good to be true"). From a general point of view, the main role of our Bayesian model is the valuation and fusion of different sources of information, respectively model- and image-based, into a single numerical formulation to be optimized. Alternative frameworks, such as fuzzy logic and evidence theory, might be worth exploring for different principles of knowledge representation and fusion. Our choice of relying on a Bayesian model is principally motivated by the significant amount of data at our disposal, enabling the derivation of meaningful statistical estimations, and by the relative simplicity and ubiquity of this framework. The field of Bayesian estimation offers particularly attractive mathematical tools such as the Monte-Carlo techniques exploited by the extraction scheme of Chapter 8.

Chapter 7

First Extraction Strategy: Minimal Path Optimization

This chapter is an extension of (Lesage et al., 2009b), to be published in Medical Image Computing and Computer Assisted Intervention (MICCAI 2009).

The first extraction strategy we propose relies on a graph-based optimization scheme inspired by popular minimal path techniques for vascular segmentation. We first present related works in the literature, exhibiting the key properties of these methods which motivated our choices. Second, we present our approach, which consists in optimizing model realizations $x^i_{1:t}$ as paths propagated on a discrete graph. In the following, we use alternatively the term *path* or *hypothesis* to designate such a realization $x^i_{1:t}$. We discuss the discretization of the search space, *i.e.*, the definition of the graph, and the key design issues of a cumulative cost metric, derived from our Bayesian formulation. One of the strong advantages of a graph-based minimal path framework is the workflow versatility it offers, from two-point interactive segmentation to fully automatic tree extraction. We finally illustrate the excellent robustness of this approach through an extensive qualitative and quantitative evaluation.

7.1 Minimal Paths for Vascular Segmentation: Background and Related Works

As was briefly discussed in Sec. 3.3.4, minimal path techniques are particularly popular for centerline-based vascular segmentation. Their core idea is to formulate centerline extraction as the optimization of a *path* given a *cumulative cost metric*. In other words, vascular centerlines are extracted as paths of minimal cumulative cost between given beginning and ending locations, referred to as the *seed* and *end* points, respectively.

7.1.1 Formulation of Minimal Path Optimization

The minimal path problem can be formulated through a cumulative energetic model on a curve $\mathcal{C}$ (Cohen and Kimmel, 1997), given by:

$$E(\mathcal{C}) = \int_\Omega (P(\mathcal{C}(s)) + w)ds = \int_\Omega (\tilde{P}(\mathcal{C}(s)))ds \tag{7.1}$$

where s is the arc-length parameter ($\|\mathcal{C}'(s)\| = 1$), Ω the support of the curve, and $\tilde{P}(\mathcal{C}(s)) = P(\mathcal{C}(s)) + w$ the so-called *potential* term. The constant w is a regulariza-

tion term, integrating to $\int_\Omega w ds = w L_\mathcal{C}$ with $L_\mathcal{C}$ the length of the curve $\mathcal{C}$. It controls the smoothness of the curve, giving, from a continuous point of view, an upper bound on the magnitude of the curvature (Cohen and Kimmel, 1997). In (Cohen and Kimmel, 1997), the authors show that this formulation corresponds to a special case of a geodesic active contour energy (Caselles et al., 1997) without a second order derivative term.

Minimal path techniques aim at optimizing, from an initial point p_0 to any point p, the *minimal action* of cumulative cost:

$$U(p) = \inf_{\mathcal{C} \in \mathcal{A}_{p_0,p}} E(\mathcal{C}) = \inf_{\mathcal{C} \in \mathcal{A}_{p_0,p}} \{ \int_\Omega (\tilde{P}(\mathcal{C}(s))) ds \} \tag{7.2}$$

which is given by the minimal cost over the set $\mathcal{A}_{p_0,p}$ of all possible paths between p_0 and p. This set is composed of paths of different lengths, defined on different domains Ω. This remark is particularly important for our application, which deals with thin, elongated structures. The computation of the minimal action $U(p)$ for all points p is referred to as the optimization of the action *map*.

The definition of a positive potential $\tilde{P}(\mathcal{C}(s))$ yields a monotonically increasing cumulative metric. This enables the *global* optimization of the action map with efficient dynamic programming schemes.

Most existing vessel-dedicated minimal path techniques follow similar algorithmic designs, iteratively optimizing the action map (Eq. 7.2). Minimal paths to the original point are extracted from the action map thanks to backtracking schemes. In practice, the action map, which theoretically covers the entire search space, is discretized. A typical approach is the use of a regular Cartesian grid over the original image.

Besides discretization strategies, techniques based on minimal paths differ in the definition of the cumulative cost. For vascular applications, the metric is designed so as to penalize paths deviating from vascular centerlines. It is customary to define potentials as the reciprocal of vessel-dedicated features, yielding high costs for low feature responses. A key characteristic of the cost metric is the *norm* employed for its numerical integration. Discrete, graph-based methods such as (Gulsun and Tek, 2008b; Wink et al., 2004) correspond to an L_1 optimization of the cost metric (the integral in Eq. 7.1 is approximated by a discrete sum). Such techniques usually employ variants of Dijkstra's algorithm (Dijkstra, 1959). L_1 optimization has generally the advantage of high computational efficiency, but can be shown to introduce metrication errors due to their discrete nature (see (Deschamps, 2001) and Sec. 7.1.3 thereafter). The fast-marching algorithm (Adalsteinsson and Sethian, 1995; Sethian, 1996a; Tsitsiklis, 1995) is an optimization scheme which can be shown to be consistent with the continuous formulation of the minimal path problem (Deschamps, 2001; Cohen and Kimmel, 1997). Fast-marching optimization approximates the true Euclidian (L_2) cumulative cost, yielding lower dependency on the discretization scheme. A counterpart, compared to discrete L_1 techniques, is an increase in computational cost due to more complex update schemes. Fast-marching schemes are particular popular for vascular segmentation (Avants and Williams, 2000; Deschamps and Cohen, 2001; Law and Heng, 2000; Lin, 2003; Wink et al., 2000b, 2001; Young et al., 2001). Finally, it is worth mentioning *fuzzy connectedness* methods (Choyke et al., 2001; Udupa et al., 2002; Yim et al., 2000, 2003), which correspond to the optimization of discrete minimal paths with respect to an L_∞ norm.

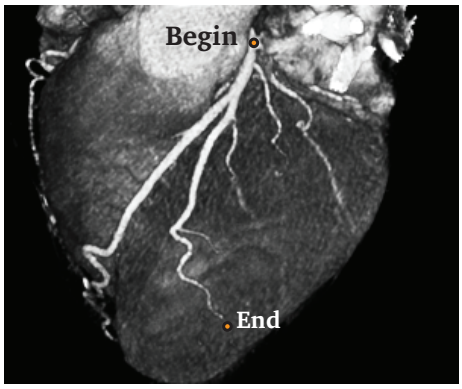


Figure 7.1: Two-point minimal path worflow. Path optimization is performed between user-provided *seed* and *end* points.

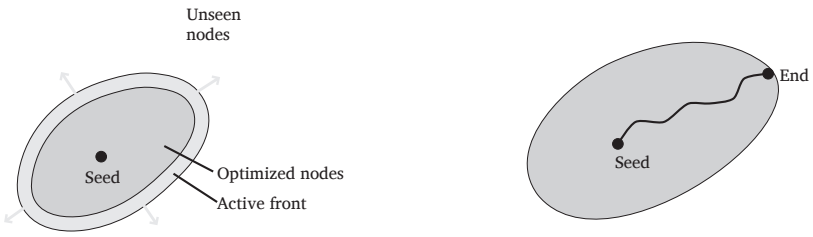


Figure 7.2: Minimal paths as a front propagation process. Left: from a seed node, graph nodes are progressively explored. Previously *unseen* nodes enter the *active front*. The set of *optimized nodes* corresponds to nodes whose optimal cumulative cost has been definitively determined. Right: from any end node in the *optimized* set, the optimal path to the seed point can be retrieved by backtracking schemes.

7.1.2 Attractive Properties

Global optimality and robustness Minimal path problems have been studied thoroughly in graph-theoretic frameworks. Classical algorithms such as Dijkstra's (Dijkstra, 1959), are known to optimize the one-seed problem (computing best paths from one seed node to all the nodes in the graph) *exactly*. The global optimality of the computed paths conveys remarkable robustness properties to such algorithms. In (Lin, 2003) for instance (see Fig. 3.22 in Sec. 3.3.4), such techniques were successfully applied to vascular centerline extraction, even in the presence of severe stenoses, occlusions or aneurysms. In (Cohen and Deschamps, 2001), minimal paths are used for *perceptual grouping*, *i.e.*, reconnecting pre-segmented vascular areas into tree structures. As will be discussed thereafter, the practical robustness of a specific minimal-path algorithm depends on several design points, among which the definition of the cost metric is key.

Computational efficiency Most minimal path techniques share similar core designs, based on an ordered propagation of optimal sub-paths. Following Dijkstra's seminal principles (Dijkstra, 1959), such designs exploit the monotonic nature of the metric to result

in computationally efficient optimization. Using a priority queue to sort partial paths, the grid is explored iteratively, focusing first on most promising sub-paths. The algorithmic complexity of both Dijkstra-like and fast-marching schemes is $O(N \log(N))$, with N the number of nodes in the graph. The computational performance is thus influenced, in particular, by the size of the graph, by the neighborhood system used, but also by the chosen norm. Although of similar algorithmic complexity, L_2 fast-marching scheme involve more intricate update schemes than L_1 or L_∞ schemes (Sethian, 1996a). The progressive exploration of the search space is often referred to as a front propagation scheme (see Fig. 7.2). In the case of discrete paths (L_1 schemes), the propagation can be considered as an ordered, multi-hypothesis and pruning strategy.

Workflow versatility A key to the popularity of minimal path techniques is the diversity of workflows in which they can be employed. For vascular applications, two-point path extraction is certainly the most classical setup. Vascular centerlines are extracted between user-provided seed and end points (Deschamps and Cohen, 2000) (see Fig. 7.1). In such scenarios, minimal paths are particularly attractive as they guarantee the control of boundary conditions. The algorithm will always find the optimal path between the specified points and will, in general, only require the exploration of a small fraction of the data to do so. Such a setup is also referred to as *partial front propagation* in the literature (Deschamps, 2001). It exploits the property that the minimal path between the seed and end points can be extracted as soon as the end point enters the set of optimized nodes (see Fig. 7.2).

More generally, minimal path techniques enable by design exhaustive search strategies. They can be used to compute the minimal paths from seeds to *all* other points by optimizing the entire action map. From a full propagation, one can then devise schemes to automatically select terminal points, or define them manually in interactive setups. Complete propagations on large search spaces can be very costly but can potentially be performed offline. Given propagated action maps, online extraction of minimal paths to any point is then extremely efficient. The latter is a particularly attractive property for interactive processing, notably for intelligent scissors/live-wire tools (Falcao et al., 2000).

Finally, minimal path frameworks can be easily adapted to accommodate for workflow refinements. For instance, one can straightforwardly allow the specification of intermediate points on the target path for increased robustness and computational efficiency, as discussed in (Lin, 2003). Minimal path optimization can thus serve as the basis for a large panel of segmentation algorithms, from highly interactive to fully automatic ones.

7.1.3 Classical Issues and Existing Refinements

Discretization issues: grid, neighborhood system and optimization norm From a continuous point of view, minimal path optimization corresponds to the extraction of geodesics given the cumulative path metric (Deschamps, 2001). The choice of grid discretization, neighborhood system and optimization norm affects the final result. In particular, Deschamps (2001) discussed *metrication* issues encountered with L_1 minimal paths. The author argued in favor of L_2 fast-marching-like optimization for its consistency, yielding better accuracy and lower dependency on the grid discretization.

Well-known undesirable effects of L_1 path optimization come principally from their discrete, step-based definition. Richer neighborhood systems (*e.g.*, 8-connectivity *versus* 4-connectivity in 2D) are advised for more accurate and more natural results. On the

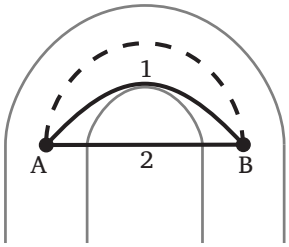


Figure 7.3: Eccentricity and shortcut issues with minimal path techniques. The ideal centerline path between points A and B is given by the dashed curve. Classical issues with minimal path techniques include eccentricity (case 1) and shortcut problems (case 2).

other hand, larger neighborhoods raise anisotropy issues, as they typically favor diagonal directions. Fortunately, as noted in (Deschamps, 2001), metrication-related problems are considerably alleviated when working on real data, given heterogeneous propagation media.

Eccentricity and shortcut issues If the target object is a thin, elongated structure such as a vessel, a frequent issue with minimal path optimization is the extraction of *eccentric* or even *shortcut* paths. As illustrated in Fig. 7.3, eccentricity issues correspond to a deviation of the extracted result from the effective centerline. Shortcuts are more severe cases where the optimized path lies outside the target vessel.

Eccentricity issues are more particularly acute when using simple intensity- or gradient-based local potentials, as in (Deschamps and Cohen, 2000). Such low-level local features do not strongly penalize deviations from the centerline location. Centering problems thus arise in large, curvy vessels, where shortest, eccentric paths within the relatively homogeneous vessel lumen are preferred to the longer, true centerline paths. In most extreme cases where the cumulative cost (and the underlying feature) is not discriminative enough against non-vascular areas, the optimal path can actually run outside the vessel lumen, resulting in so-called shortcut effects. Heuristic approaches have been proposed to reduce eccentricity issues, relying for instance on distance-based recentering techniques from a binary contour segmentation (Deschamps and Cohen, 2000), coupled with a non-cumulative fast-marching variant in (Wan et al., 2002a). Another approach to reduce shortcut issues when dealing with elongated structures is the so-called *recursive* fast-marching modification proposed in (Deschamps, 2001; Deschamps et al., 1999). The idea is to alter the update scheme to favor low cost paths further. The modification introduced is inherently heuristic as it violates the monotonic nature of the cost optimization. However, it is interesting to note that the motivation behind this scheme is closely related to the ideas we discuss in Sec. 7.2.3. From a general point of view, eccentricity and shortcut issues may have various origins. Discretization-related metrication errors we previously mentioned is one of them. We believe however that these practical problems arise principally from the cumulative nature of the optimization, as discussed more in depth in Sec. 7.2. A key factor in limiting them is designing an adequate cost metric. This task is particularly delicate for elongated, complex structures such as vessels. First, it is advisable to design the metric around a vessel-dedicated medialness feature yielding higher responses at centerline

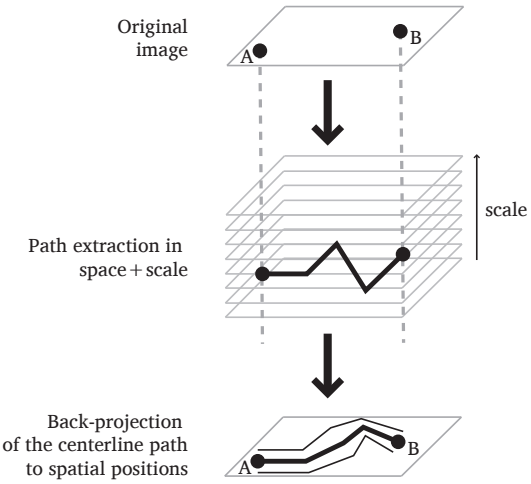


Figure 7.4: Multiscale minimal path methods such as (Wink et al., 2004; Li and Yezzi, 2007) add an explicit scale dimension to the optimization graph. The result centerline path is obtained straightforwardly by retaining only the spatial components of the optimal space+scale path, which corresponds to a projection operation. The associated scale information can be used to derive an approximate contour segmentation.

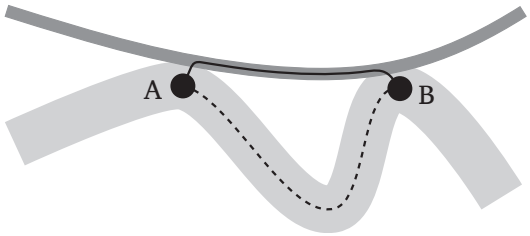


Figure 7.5: Disambiguation with scale information. With classical minimal path methods for centerline extraction, the presence of a different scale, straighter and/or more contrasted vessel near the target vessel may result in the extraction of an erroneous path (black plain curve). This problem is particularly critical for 2D projection acquisitions such as X-ray angiography, as different vessels may overlap. As illustrated through different experiments in (Wink et al., 2004), the explicit incorporation of scale information in the minimal path optimization can help disambiguate such situations in order to extract the desired centerline (dashed curve). Illustration based on material from (Wink et al., 2004).

locations while strongly penalizing non-vascular structures. Hessian-based features such as the vesselness from (Frangi et al., 1998) were for instance applied in minimal path frameworks in (Metz et al., 2008a; Olabarriaga et al., 2003; Wink et al., 2001), while a ray-based oriented medialness measure was preferred in (Gulsun and Tek, 2008b). Robustness to shortcut and eccentricity problems is strongly correlated with the discriminative power of the underlying image feature.

Multiscale minimal paths The aforementioned vascular-dedicated minimal path methods (Metz et al., 2008a; Olabarriaga et al., 2003; Wink et al., 2001; Gulsun and Tek, 2008b) rely on multiscale features but retain classical, spatial centerline-based designs. The optimization grid corresponds to spatial centerline positions. Multiscale features are used to valuate spatial potentials, typically by selecting the maximal feature response over scales at a given location. This corresponds to a projection of the scale space feature landscape onto the 2D or 3D geometric space *before* path extraction. Although the maximally responding scale can be recorded to obtain a rough surface segmentation (Gulsun and Tek, 2008b), scale information is not taken into account into the optimization scheme. Alternatively, recent works have proposed to incorporate the scale space directly in the path optimization process as an additional dimension (Wink et al., 2004; Li and Yezzi, 2007; Benmansour et al., 2009). As schematically depicted in Fig. 7.4, the optimized scale+space paths are simply projected onto spatial locations to recover the centerline curve. The associated scale information can be directly exploited as an approximate contour segmentation. In practice, such multiscale minimal paths techniques have several advantages. As illustrated in Fig. 7.5, they help disambiguating scale-related issues by favoring smooth, coherent scale transitions. As noted in (Wink et al., 2004), this benefit is arguably greater in 2D projection images than in 3D, where vessels should not overlap. Nonetheless, taking scale information into account in the optimization process yields a significant gain in robustness, notably resulting in better centered paths (Li and Yezzi, 2007).

A drawback of multiscale minimal paths is that the total size of the graph to be explored is multiplied with the number of considered scales. This significantly increases costs in memory and computational time. A second general issue is the intricate design of an adequate multiscale metric. Existing multiscale minimal path methods introduce extra parameters, to control for instance the weighting between space and scale transitions (Li and Yezzi, 2007; Wink et al., 2004; Benmansour et al., 2009; Pechaud et al., 2009). The method detailed in this chapter is directly inspired by these recent developments on multiscale minimal paths and proposes some refinements to alleviate their principal drawbacks, namely their computational and memory cost, along with the intricate definition of an adequate cumulative cost metric.

Anisotropic/directional propagation Another recent trend of interest is the development of *anisotropic flows* for vascular-dedicated L_2 fast-marching methods (Lin, 2003; Benmansour et al., 2009; Benmansour and Cohen, 2009). Anisotropic propagation flows are designed to favor paths in certain directions, typically based on local estimations of the principal vessel direction. As thoroughly discussed in (Lin, 2003), anisotropic propagation yields more robust results, decreasing the sensitivity to shortcut effects. Lin (2003) additionally introduced a heuristic *tip-tracking* modification of anisotropic fast-marching, further favoring the orientation given by the tip of the propagation front. In (Benmansour et al., 2009), the authors coupled an anisotropic flow with the idea of multiscale path optimization previously mentioned. Their method was later successfully applied to 3D CTA coronary segmentation in (Benmansour and Cohen, 2009). Benmansour et al. (2009) thoroughly discussed the construction of an adequate multiscale anisotropic metric, resulting in a homogeneous and elegant framework. The main drawback of their approach is the relative complexity of the update schemes it involves. Despite the use of more efficient numerical approximations (Benmansour et al., 2009; Benmansour and Cohen, 2009), such a method remains relatively costly in 3D.

Directional information can also be taken into account by considering the orienta-

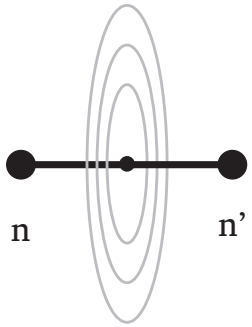


Figure 7.6: The oriented multiscale medialness feature from (Gulsun and Tek, 2008b) is computed along edges between neighboring graph nodes n and n' to valuate transition costs in a minimal cost path framework.

tion domain as an additional dimension for the path optimization. Such an approach was first employed for robot navigation in (Kimmel and Sethian, 2001). Deschamps (2001) later discussed its potential for medical applications, notably for the directional regularization of virtual endoscopy. The idea of considering extra dimensions (in addition to spatial ones) is shared with the multiscale path techniques we previously mentioned. Very recently, Pechaud et al. (2009) demonstrated the extraction of 4D paths (position+scale+orientation) on 2D road networks and angiograms. The use of the additional orientation dimension helps considerably the disambiguation of overlap situations encountered on 2D angiographies. The method of Pechaud et al. (2009) introduces parameters to control spatial versus angular variations, which closely relate the anisotropic ratio found in the work of (Benmansour et al., 2009). Finally, the principal drawback of this approach is the combinatorial growth of the search space yielded by the additional dimension. We believe the extension of this method to 3D would be prohibitively costly, time- and memory-wise, for current standard workstations.

A last, straightforward way of integrating direction-dependent path valuation was first hinted at in (Wink et al., 2004) and recently demonstrated in (Gulsun and Tek, 2008b) with L_1 path optimization for vascular centerline extraction. As illustrated in Fig. 7.6, Gulsun and Tek (2008b) valuate edge costs given a multiscale feature *oriented* by the edge direction. In other words, the graph edges serve as a discretization of possible orientations. This scheme avoids the introduction of additional dimensions by exploiting the (admittedly coarse) directional information already embedded in the spatial graph. We reuse this idea in our approach.

7.2 From the Bayesian Model to an Empirical Cumulative Cost Metric

One of the key design issues of minimal path methods is the definition of an adequate cumulative cost metric. We now discuss our proposal for such a metric, which we derive from our recursive Bayesian model.

7.2.1 Cumulative Formulation of the Bayesian Model

We recall that our geometric model (Chapter 3.1.2) is defined as a discrete series $x_{1:L}$ of states $x_t = (p_t, d_t, r_t)$, with centerline positions p_t, directions d_t and radius r_t. Centerline positions and tangential directions are coupled thanks to control points c_t, so that a state can alternatively be defined by successive control points $x_t = (p_t, d_t, r_t) \equiv (c_{t-1}, c_t, r_t)$. In Chapter 6, we discussed our Bayesian formulation, notably through the expression of the posterior probability of the model given sets of data observations $z_t = \{y_t^i\}$. Observations z_t available at time correspond to elementary observations y_t^i linked to state realizations x_t^i.

It is possible to link our seminal Bayesian formulation to minimal path problems by noting that maximizing the posterior probability is equivalent to minimizing its negative logarithm. A similar approach, reformulating a Bayesian model for minimal path optimization, was notably proposed in (Rousson et al., 2006) for the segmentation of the esophagus from CT data. The MAP problem with boundary constraints, between the initial state x_0 and the final state x_f, could be formulated as finding the realization $x^*_{1:L^*}$ as:

$$x^*_{1:L^*} = \arg \min_{\mathcal{X}_{x_0,x_f}} \left(-\log(p(x_{0:L}|z_{1:L})) \right) \tag{7.3}$$

Again, an important remark is that the optimal length L^* minimizing the posterior between x_0 and x_f is not known. The set $\mathcal{X}_{x_0,x_f}$ of possible realizations with initial state x_0 and final state x_f comprises candidates of different lengths. Comparing such realizations through their posterior probabilities is inherently ambiguous, as these probabilities are normalized on different definition domains depending on their respective length L.

Developing the negative log-posterior $-\log(p(x^i_{0:t}|z^i_{1:t})$ for our particular model (Eq. 6.5, 6.11, 6.10), we obtain:

$$\begin{aligned}
-\log(p(x^i_{0:t}|z_{1:t})) &= -\log(p(x^i_{0:t-1}|z_{1:t-1})) - \log(p(x^i_t|x^i_{t-1})) - \log(p(z_t|x^i_t)) \\
&\quad + \log(p(z_t|z_{1:t-1})) \\
&= -\log(p(x^i_{0:t-1}|z_{1:t-1}) - \log(p(r^i_t|r^i_{t-1})) - \log(p(d^i_t|d^i_{t-1}, r^i_{t-1})) \\
&\quad - \log\left(\frac{p_v(y^i_t|r^i_t)}{p_{bg}(y^i_t)}\right) - \sum_{y^j_t \in z_t, \forall j} \log(p_{bg}(y^j_t)) + \log(p(z_t|z_{1:t-1}))
\end{aligned} \tag{7.4}$$

One can note that the negative log-posterior of a realization is expressed in an additive, recursive form, and is strictly increasing with its length t. Theoretically, it could thus be a potential candidate for minimal-path-like optimization. The development of Eq. 7.4 poses however a few fundamental problems:

- certain terms, such as $\log(p(z_t|z_{1:t-1}))$ and the sum of background likelihoods $\sum_{y^j_t \in z_t, \forall j} \log(p_{bg}(y^j_t))$, cannot be computed analytically and should be integrated over all possible states, which is impractical in minimal path frameworks;
- the directional prior $\log(p(d^i_t|d^i_{t-1}, r^i_{t-1}))$ raises discretization issues, discussed in Sec. 7.3.

Our formulation relates closely to that of recent works on multiscale minimal paths for vascular segmentation (Li and Yezzi, 2007; Benmansour et al., 2009). In (Benmansour et al., 2009) in particular, the authors introduce parameters to control radius variations with respect to spatial ones. In contrast, we use an explicit prior on such variations, and both prior and likelihood terms are learned automatically from our database.

7.2.2 Empirical Cumulative Cost

The empirical cost metric we propose is a simplification of the log-posterior in Eq. 7.4. We define the cost $E(x^i_{0:t})$ of a realization $x^i_{0:t}$ in a recursive fashion:

$$E(x^i_{0:t}) = E(x^i_{0:t-1}) + C(x^i_{t-1}, x^i_t) \tag{7.5}$$

where $C(x^i_{t-1}, x^i_t)$ is the *update* cost from state x^i_{t-1} to state x^i_t. We define it as:

$$C(x^i_{t-1}, x^i_t) = -\log(p(r^i_t|r^i_{t-1})) - \log\left(\frac{p_v(y^i_t|r^i_t)}{p_{bg}(y^i_t)}\right) + M \tag{7.6}$$

where we fix M as a positive constant, as discussed in the next section. Compared to Eq. 7.4, our update cost formulation retains only the terms depending on the realization $x^i_{0:t}$ at hand. One can also note that the directional prior is left out, for reasons discussed in Sec. 7.3.

The cost of a realization/hypothesis/path is valued cumulatively with the addition of a new state x^i_t to $x^i_{0:t-1}$. Intuitively, the update cost $C(x^i_{t-1}, x^i_t)$ depends only on the preceding state x^i_{t-1} (1^{st} order Markovian model) and is composed of two main parts:

- a model-based prior on radius transitions: $-\log(p(r^i_t|r^i_{t-1}))$;
- a data attachment term, the logarithm of the likelihood ratio: $-\log\left(\frac{p_v(y^i_t|r^i_t)}{p_{bg}(y^i_t)}\right)$.

7.2.3 Fixing the Constant M

To be able to optimize our cumulative cost (Eq. 7.5) in an efficient, dynamic-programming-like manner, it should be increasing. This translates as a constraint on the update cost being positive. With respect to the constant M, we obtain:

$$\begin{aligned} C(x^i_{t-1}, x^i_t) &\geq 0 \\ -\log(p(r^i_t|r^i_{t-1})) - \log\left(\frac{p_v(y^i_t|r^i_t)}{p_{bg}(y^i_t)}\right) + M &\geq 0 \\ M \geq \log(p(r^i_t|r^i_{t-1})) + \log(p_v(y^i_t|r^i_t)) - \log(p_{bg}(y^i_t)) & \end{aligned} \tag{7.7}$$

The constant M plays a role similar to the regularization term w in the continuous formulation of minimal paths (Eq. 7.1). On a side note, it is interesting to note that the terms we omitted from the true log-posterior (Eq. 7.4) correspond globally to the normalization of the probabilities. They can alternatively be interpreted as a *varying* regularization term, which depends on the generation t [1].

As was previously mentioned, one key issue with the minimal path formulation is the comparison of paths of potentially widely varying lengths, which translates into considerations on the normalization of the incremental costs and the regularization of the extracted paths. The use of a positive, increasing cumulative cost metric inherently favors paths of small lengths. The introduction of additional regularization ($w > 0$ in Eq. 7.1), although often desirable for the smoothness of the result, further amplifies the penalization of long

1. The term $p(z_t|z_{1:t-1}) = \int p(z_t|x_t)p(x_t|z_{1:t-1})dx_t$, integrated over all states of generation t, can be interpreted as a measure of agreement between the marginal likelihood and the prediction $p(x_t|z_{1:t-1})$. It can finally be viewed as the scaling of individual likelihood values $p(y^i_t|x^i_t)$ with respect to the overall data innovation.

realizations. The prevalence of small paths has particularly strong implications for applications dealing with curved, elongated structures such as vessels. As was discussed in Sec. 7.1.3, such applications are particularly prone to eccentricity and shortcut issues. This observation led some authors to go as far as heuristically breaking the cumulative properties of the optimization, proposing for instance so-called *non-cumulative* (Wan et al., 2002b) and *recursive* variants (Deschamps, 2001; Deschamps et al., 1999) of the fast-marching algorithm.

Our empirical approach is to set the constant M to the lowest possible value satisfying Eq. 7.7. Numerical issues arise because of the third term, $-\log(p_{bg}(y_t^i))$. In practice, we impose a lower bound $p_{min} = 10^{-6}$ on all probability values. We thus obtain:

$$M = \max_{r_{t-1},r_t} \log(p(r_t|r_{t-1})) + \max_{r_t,y_t} \log(p_v(y_t|r_t)) - p_{min} \tag{7.8}$$

It is interesting to reinject that value into the expression of the update cost (Eq. 7.6). We now have:

$$C(x_{t-1}^i, x_t^i) = -\log\left(\frac{p(r_t^i|r_{t-1}^i)}{\max_{r_{t-1},r_t} p(r_t|r_{t-1})}\right) - \log\left(\frac{p_v(y_t^i|r_t^i)}{\max_{r_t,y_t} p_v(y_t|r_t)}\right) - \log\left(\frac{p_{min}}{p_{bg}(y_t^i)}\right) \tag{7.9}$$

One can see that the minimal update cost C_{min} is now exactly zero. This implies that a "perfect" realization (chaining at every step the best radius transition, highest vessel likelihood and lowest background likelihood) will retain a cumulative of zero regardless of its length. In effect, the penalization on long paths is as low as possible while still having a positive update. This choice, coupled with the discriminative power of our likelihood model (see Sec. 6.3 and Fig. 6.4), contributes to the robustness of our approach to shortcut effects.

7.2.4 Discussion

Our empirical cumulative cost function is derived from our Bayesian model, from which we omit intractable normalization terms. In addition, we introduced further adjustment (through the constant M) to limit potential shortcut effects. We emphasize the empirical nature of this metric, which sensibly diverges from our seminal Bayesian formulation. Our approach does not cope with the original maximum *a posteriori* formulation (Eq. 7.3), but in return benefits from a particularly efficient and flexible algorithmic framework. In fact, if one omits the directional prior, the cumulative cost $E(x_{0:t}^i)$ of a realization $x_{0:t}^i$ relates to the negative log-posterior probability as follows:

$$E(x_{0:t}^i) = -\log(p(x_{0:t}^i|z_t)) + N(t)$$

where $N(t)$ is a normalization constant dependent on the length t of the realization. In effect, the comparison of two realizations of same length still corresponds to a maximum *a posteriori* selection rule. Our choice of M can be interpreted as an empirical alteration of the normalization scheme. This is particularly clear for the radius prior and vessel likelihood terms in Eq. 7.9, which can be viewed as scaled with respect to their respective maximums.

The discrete nature of our geometric and Bayesian model led to an additive cost function, expressed as a discrete sum over successive states of model realizations. We propose to optimize the cost of model realizations propagated on a discrete graph. This corresponds to an L_1 minimal path scheme, which, as discussed in the previous sections, is dependent on the underlying discretization of the search space.

7.3 Graph and Neighborhood Structures

In this section, we discuss the discretization of the search space engendered by our geometric model. Our goal is to define a graph and associated neighborhood structures suitable for minimal path optimization.

States $x_t = (p_t, d_t, r_t)$ are composed of three components, namely the centerline position p_t, tangential direction d_t and radius r_t. A particularity of our geometric model (Chapter 4) is to link positions and tangential directions through the introduction of *control points* c_t. States are then equivalently given by $x_t = (p_t, d_t, r_t) \equiv (c_{t-1}, c_t, r_t)$, reducing the parameterization to 4 dimensions, 3D control points $\{c_t\}$ and radius values. The graph we employ is thus a 4D structure, defined in two parts:

- a 3D grid discretizing control point locations;
- a range of discretized radiuses.

7.3.1 Discretization of Control Point Locations

As was mentioned in Sec. 7.1.3, L_1 minimal path optimization is dependent on the underlying discretization. The choice of the grid and associated neighborhood system is a tradeoff between computational efficiency and accuracy. Impacting factors are the richness of the neighborhood, its spatial and angular isotropy and symmetry properties.

In order to discretize the possible locations of successive control points, it is classical to use a regular Cartesian grid with 6- or 26-neighborhood. As an alternative to Cartesian grids, we opted for a Body-Centered Cubic (BCC) grid and its associated 14-neighborhood system (see Fig. 7.7). The BCC grid is the optimal 3D sampling grid with respect to Shannon's theorem and is one of the most compact (in terms of sphere packing) (Conway and Sloane, 1999). Combined with its strong symmetry properties, this makes the BCC grid particularly appealing, for instance for discrete mathematical morphology (Serra, 1983). Whereas a classical 26-neighborhood is highly anisotropic (with a factor of $\frac{\sqrt{3}}{1} = 1.73$ between closest and most distant neighbors), the BCC 14-neighborhood enjoys an anisotropy factor of only $\frac{1}{0.5\sqrt{3}} = 1.155$ between closest and most distant neighbors (see Fig. 7.7). On the other hand, it is more limited in terms of direction discretization, an issue fortunately alleviated by the tolerance of the `MFlux` feature with respect to orientation changes (see Chapter 5). Finally, it remains simple to encode and efficient to access. As depicted in Fig. 7.7, a BCC grid can be obtained by interlacing two shifted Cartesian grids (Matej and Lewitt, 1995).

A more formal study would be necessary to precisely quantify the advantage of the BCC grid compared to 6- and 26-connected Cartesian grids. In our experiments, we found 6-connectivity to be insufficient, resulting in jagged and inaccurate centerline paths. Results obtained with 26-connectivity were more satisfying. The BCC grid slightly outperformed the Cartesian, 26-connected grid in terms of robustness and accuracy. In particular, the BCC implementation seemed less prone to local shortcut effects. We attribute this result to the reduction of metrication errors due to the lower anisotropy of the neighborhood.

For our coronary application, we use an isotropic BCC grid with the same intra-slice resolution as the original data (typically 0.3mm per voxel). The inter-slice grid spacing (between $z = 0$ and $z = 0.5$ on Fig. 7.7) is thus set to half the intra-slice resolution of the data. Finally, one can note that possible *centerline* points $\{p_t\}$ lie *in between* grid nodes, following the relation $p_t = \frac{c_t + c_{t-1}}{2}$. They correspond to the centers of the edges defined by the 14-connected BCC neighborhood. These edges also define the possible directions

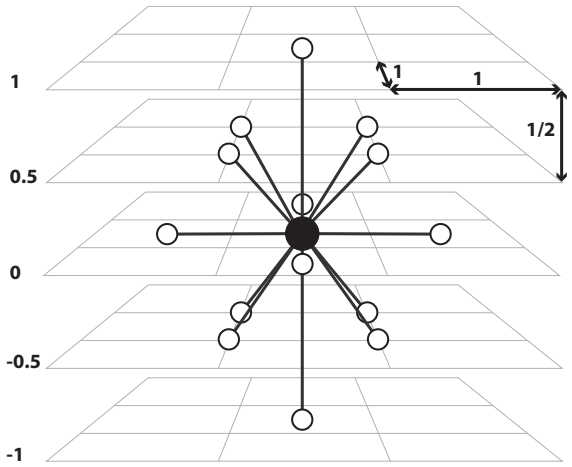


Figure 7.7: Body-centered cubic (BCC) 14-neighborhood. White spheres are the 14 neighbors of the center position (in black). Note the grid shift of interleaved slices and the spacing indicated by arrows. Fixing the center coordinates to $(x_0, y_0, z_0) = (0, 0, 0)$, the 14-BCC neighbors are: $(0, 0, 1)$, $(-0.5, -0.5, 0.5)$, $(0.5, -0.5, 0.5)$, $(-0.5, 0.5, 0.5)$, $(0.5, 0.5, 0.5)$, $(0, -1, 0)$, $(-1, 0, 0)$, $(1, 0, 0)$, $(0, 1, 0)$, $(-0.5, -0.5, -0.5)$, $(0.5, -0.5, -0.5)$, $(-0.5, 0.5, -0.5)$, $(0.5, 0.5, -0.5)$, $(0, 0, -1)$.

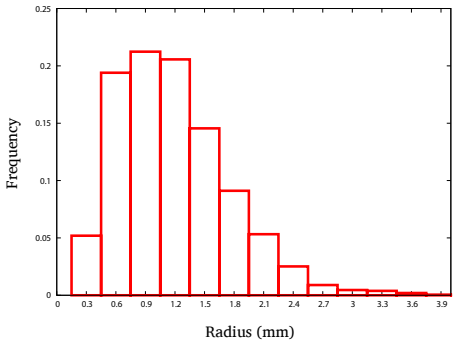


Figure 7.8: Frequency histogram of radiuses in our database (858 coronary branches, 61 datasets).

$\{d_t\}$ as $d_t = \frac{c_t - c_{t-1}}{\|c_t - c_{t-1}\|}$.

7.3.2 Discretization of Possible Radiuses

The definition of the spatial grid for control point locations fixes only part of our parameter space, namely the positions of centerline points and tangential directions. We now discuss the discretization of possible radiuses.

In vascular applications, radiuses can be considered as a scale parameter. For applications potentially dealing with large radius ranges of vascular structures, it is customary, especially for filtering, to use an exponential discretization of the candidate radius values

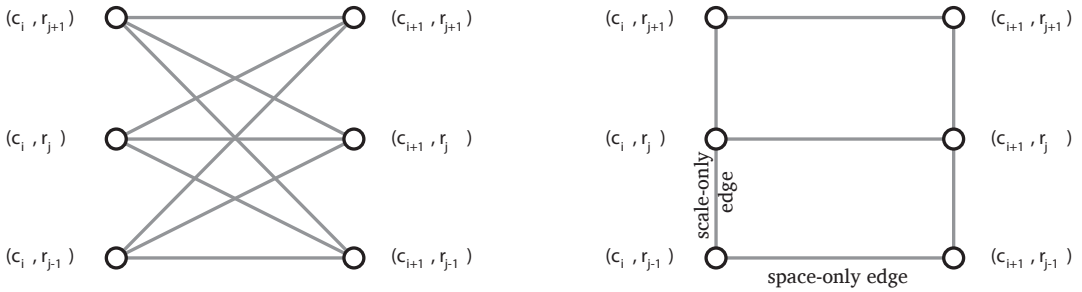


Figure 7.9: Definition of the space+scale 4D graph. A partial graph (two spatial locations, three scales/radiuses) is illustrated. Left: the graph engendered by our neighborhood system. Right: the graph used in (Li and Yezzi, 2007; Benmansour et al., 2009).

of the form $r^i = r_{min} \times (f_r)^i$, with f_r a factor slightly greater than 1 (see, among others, (Krissian et al., 2000)). The motivation is to keep a relatively low number of possible radiuses while still covering the entire predefined range. Such schemes consider that smaller vessels require greater accuracy to be captured, while lower accuracy is tolerable for bigger vessels. In contrast, for our particular application, the maximal coronary radius is small (compared to the aorta, for instance), yielding a relatively narrow scale range. Fig. 7.8 gives the frequencies of coronary radiuses encountered on our ground-truth database. In both the extraction schemes we propose, we discretize possible coronary radiuses between $r_{min} = 0.3mm$ and $r_{max} = 3.9mm$ uniformly, by steps of $s_r = 0.3mm$, resulting in a set R of 13 different radius values.

We chose a uniform discretization so as not to favor smaller coronaries over bigger ones. In order to chose a suitable discretization step, it is important to remark that the relative scale dynamics of coronary arteries are important, from a factor of 1 to 13 between r_{min} and r_{max}. In particular, we want to be able to capture typical scale variations along the artery (see Sec. 6.4.1). A discretization step of $s_r = 0.3mm$ (of the order of the typical intra-slice resolution of the data) proved to be sufficient for this purpose (see Sec. 7.3.5). This choice is a trade-off between accuracy and computational efficiency. Over several sets of experiments, we did not notice significant improvements when decreasing s_r, neither on robustness nor accuracy criteria. In particular, the radius estimation error did not noticeably decrease for smaller values, leading us to conclude that radius discretization was not the bottleneck factor in this regard.

Finally, in terms of possible radius transitions, we consider for the moment that the potential successors $\{r_t\}$ of a radius r_{t-1} are *all* radiuses in $R = [\![r_{min}, r_{max}]\!]$. This defines the radius-related part of our neighborhood system. In Sec. 7.4, we will discuss the possible restriction *a priori* of this neighborhood.

7.3.3 Definition of the 4D Graph

We obtain our complete space+scale neighborhood as the product set of spatial and radius neighborhoods. The search space induced by our space+scale neighborhood is thus the oriented 4D graph $G = \{V, E\}$, with vertices/nodes $V = \{(c_i, r_j)\}$ corresponding to couples of control point locations (nodes of the 3D spatial grid) and discrete radiuses. We emphasize that the states $x_t = (p_t, r_t, d_t) \equiv (c_{t-1}, c_t, r_t)$ are implicitly encoded as the *edges*

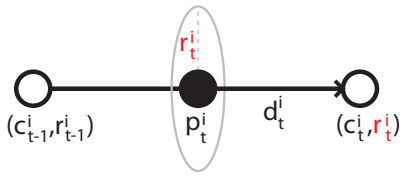


Figure 7.10: 4D graph and model states. Depicted as white circles are two neighboring nodes (c^i_{t-1}, r^i_{t-1}) and (c^i_t, r^i_t) of the 4D graph. Two such consecutive nodes fully determine the state $x^i_t = (p^i_t, d^i_t, r^i_t) \equiv (c^i_{t-1}, c^i_t, r^i_t)$, given our geometric model (Chapter 4). The position p^i_t is given by $p^i_t = \frac{c^i_t + c^i_{t-1}}{2}$ (center of the spatial edge $c^i_t - c^i_{t-1}$). The direction $d^i_t = \frac{c^i_t - c^i_{t-1}}{\|c^i_t - c^i_{t-1}\|}$ is that of the same edge. Finally, the radius r^i_t is directly given by the second 4D node (c^i_t, r^i_t).

of the 4D graph (see Fig. 7.10). This allows us to take into account coarse directional information without augmenting the dimension of the graph.

Closely related multiscale graphs are also found in (Wink et al., 2004; Li and Yezzi, 2007; Benmansour et al., 2009), but with different structures and different optimization contexts. In particular, the methods of Li and Yezzi (2007) and Benmansour et al. (2009) develop fast-marching-based L_2 multiscale flows. Their 4D scale+space neighborhoods consist of space-only or scale-only neighbors, as illustrated in Fig. 7.9, right. In a L_2 context, the scale+space neighbors serve as anchors for the interpolation/update scheme (the so-called upwind scheme). The more neighbors are taken into account, the more accurate the computation is. The optimal path is not constrained to lie on the grid and is sub-voxel accurate. Finally, these techniques employ node-based *potentials* to define their L_2 cost metrics. In our L_1 discrete framework, the scale+space neighborhood defines the discrete set of all possible transitions for our model (Fig. 7.9, left). Our metric values these transitions, *i.e.*, the graph *edges*. In Sec. 7.4, we introduce different heuristics to alleviate the computational burden imposed by such large neighborhoods.

7.3.4 On the Use of Directional Information

Directional information can either come from the image (local estimation) or from a model. We previously mentioned anisotropic fast-marching methods (Lin, 2003; Benmansour et al., 2009; Benmansour and Cohen, 2009) as examples of works exploiting data-derived orientation cues. In such cases, orientation estimations are local and depend solely on the image data. The use of model-based directional information raises different theoretical and practical issues. In L_2 optimization framework, we mentioned the possibility of imposing an upper bound on the curvature radius through the regularization constant (see (Cohen and Kimmel, 1997) and Sec. 7.1.1). In this work, we initially considered the integration of our directional prior $p(d_t|d_{t-1})$ (Sec. 6.4.2) into the minimal path computation. Our directional prior involves three consecutive control points, which translates, in terms of graph-based optimization, into second order constraints. In other

words, the valuation of an edge cost would then depend on the preceding edge[2]. In order to account for second-order constraints and still guarantee the validity of a Dijkstra-like optimization scheme, a possibility would be to consider an extended graph, where a single edge corresponds to two consecutive edges of the initial graph. In effect, this could be carried out by the introduction of additional, orientation-related dimensions, similarly to (Pechaud et al., 2009). Such a solution would cause a combinatorial growth of the search space, which we explicitly tried to avoid by coupling spatial and directional dimensions in our geometric model. It would also raise a number of discretization-related problems, since a relatively fine discretization of possible orientations would be required to adequately represent our directional prior. For the sake of computational efficiency, we eventually decided against the use of such techniques and left out the directional prior. One can note however that the data likelihood is extracted in an oriented fashion, as discussed in the following paragraph.

7.3.5 Edge Valuation

The cumulative update $C(x_{t-1}^i, x_t^i)$ in Eq. 7.9 corresponds to a cost on an edge of the 4D graph. Although it would be theoretically possible to valuate the weights of all edges and explicitly construct the full graph, this would be particularly inefficient. Our algorithm generally explores only a small subset of the entire graph to extract the target vessel(s). We thus compute update costs dynamically, when they are required.

The first term in Eq. 7.9 corresponds to the log-prior on radius transitions. As was previously discussed, radius realizations are discretized in a range $[r_{min}, r_{max}]$, sampled uniformly by steps of $s_r = 0.3mm$. This resolution was deemed sufficient to capture the main radius variations embedded in the joint density $\hat{f}(r_t|r_{t-1})$ learned in Sec. 6.4.1. This discretization requires however the numerical reintegration of the prior PDFs $p(r_t|r_{t-1})$ from the continuous joint density, following Eq. 6.6. We obtain histogram representations, as depicted in Fig. 7.11.

The computation of the update cost also involves the valuation of the log-likelihood terms. Similarly to the radius transition prior, the PDFs $p_v(y_t|r_t = r_t^i)$ are numerically reintegrated from the continuous joint density (Fig. 6.2) given the discretized radius values. The observation y_t^i associated to the state $x_t^i = (p_t^i, r_t^i, d_t^i)$ is obtained as the `MFlux` response for corresponding state parameters (see Fig. 7.10). In practice, we thus use the edges of the spatial grid to locate and orientate `MFlux` computations. This scheme is closely related to (Gulsun and Tek, 2008b) (see Fig. 7.6).

7.4 Restriction of Scale-Space Exploration

Theoretically, one could directly apply Dijkstra's algorithm on the full 4D graph and obtain a globally optimal path. The biggest issue with such an approach is the size of the graph. The spatial grid can reach a size of the order of 512^3 nodes, multiplied by $|R| = 13$ radiuses. Such graph sizes raise coupled problems of memory and computational cost on today's standard workstations.

Our algorithm propagates hypotheses (realizations of our model) as paths on the 4D space+scale graph. As a first mechanism to restrain the potential combinatorial explosion

2. One can note that the scale prior is a first order term, valuated on a single edge of the space+space graph.

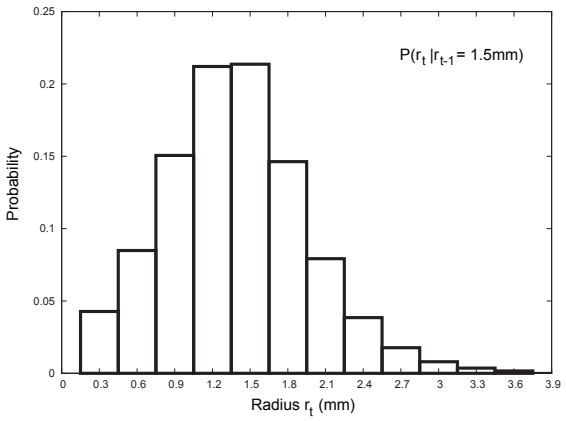


Figure 7.11: Sample of discretized radius prior PDF used in practice.

of the exploration, we make use of the well-spread principles governing dynamic programming/minimal path methods. In that sense, our approach is closely related to the recent developments in multiscale minimal path techniques discussed previously. A contribution of our implementation is the introduction of different strategies limiting scale-space exploration for increased computational and memory efficiency.

7.4.1 Static, *A Priori* Pruning of Scale Transitions

A first strategy to reduce the computational burden of our algorithm is to restrict the set of radiuses $\{r_t\}$ considered as potential successors of r_{t-1}. This can be done by evaluating only the N_r most probable radius transitions *a priori*, in the sense of $p(r_t|r_{t-1})$ (see Sec. 6.4.1 and Fig. 7.11). In other words, hypotheses with the less probable, *i.e.* strongest radius variations are pruned *a priori*. This restriction of possible radius variations is heuristic by nature and changes the search space, as it actually removes edges from the 4D graph. In our tests however, we observed virtually no impact on the extracted result for as low as $N_r = 4$. In the example of Fig. 7.11, with $r_t = 1.5mm$, this implies that we now consider only neighboring states $\{x_t\}$ with $r_t = 0.9, 1.2, 1.5$ or $1.8mm$. The histogram distributions are re-normalized accordingly.

7.4.2 Dynamic Selection of Propagated Hypotheses

Scale-space exploration can also be constrained *dynamically*, following a simple observation: in practice, only a small fraction of the scale-space, around the local radius of interest, is relevant. Aberrant radius realizations are effectively penalized by likelihood terms in Eq. 7.9 and are thus likely to be considered relatively late in the ordered exploration process. Nevertheless, a classical implementation of the full multiscale path optimization should account for the *possibility* of reaching them. The induced cost is particularly high in terms of memory requirements. We propose to explicitly limit the amount of scale-space exploration by locally and dynamically restricting the set of radiuses considered.

Many path realizations can share a spatial location c_i. In the multiscale approach, hypotheses reaching c_i with different radiuses r_j and r_k are linked to different nodes of the 4D graph, respectively (c_i, r_j) and (c_i, r_k). On the other hand, among the hypotheses reaching a node (c_i, r_j), *i.e.* hypotheses with the same radius r_j at location c_i, it is

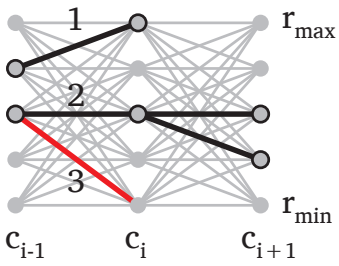


Figure 7.12: Restriction of the scale-space exploration. Our algorithm propagates only the H best hypotheses of different scale values reaching a spatial location. Here, with $H = 2$, selected hypotheses are highlighted in bold. See text for details.

only necessary to retain the best one. This observation is of importance for the exact formulation of our approach:

> We propagate only the H *best* hypotheses of *different* scales reaching a given spatial location.

We use a fixed and relatively low number $H < |R|$, *e.g.* up to 4 hypotheses per location c_i. As depicted in Fig. 7.12, the set of selected radiuses varies dynamically, in function of the parent hypotheses reaching a location c_i. For instance, hypotheses 1 and 2 are the two selected ones for location c_i. Hypothesis 3 is pruned. From c_i to c_{i+1}, hypothesis 1 does not engender any selected child, while hypothesis 2 manages to propagate two of its children. As the local selection is based on the previous partial paths, this heuristic does not guarantee the global optimality of the extracted paths. However, it focuses efficiently on the locally most relevant radiuses, leading to significant computational and memory gains compared to full 4D exploration. In practice, this makes our method usable on today's commercial workstations.

Our Dijkstra-like propagation algorithm is discussed in details in the following sections. Compared to textbook implementations, it induces a slight computational overhead linked to the maintenance of the H selected hypotheses per control location. For $H = |R|$, our method would be equivalent to a full 4D graph exploration. The global optimality of the result would be guaranteed and the additional burden of sorting the H hypotheses per location would become unnecessary. Conversely, for $H = 1$, our algorithm becomes similar, although not exactly equivalent to a classical spatial path optimization scheme such as (Gulsun and Tek, 2008b). Gulsun and Tek (2008b) rely on a multiscale medialness feature to value the graph edges, simply selecting the maximal response over scales. In other words, scale selection is performed as a projection, independently for each spatial location. In contrast, we use a prior on radius transitions and emphasize that even for $H = 1$, our approach still involves a dynamic, path-dependent radius selection process. In particular, our algorithm will not necessarily span the entire radius range for each edge, depending on the allowed transitions from the previous neighbors. More generally, varying $1 < H < |R|$ controls the tradeoff between computational efficiency and robustness to scale-related ambiguities.

Finally, one should note that our approach solely restricts scale-space, not spatial exploration. In particular, we still guarantee the possibility of extracting a path between any two spatial points, which is critical for interactive workflows.

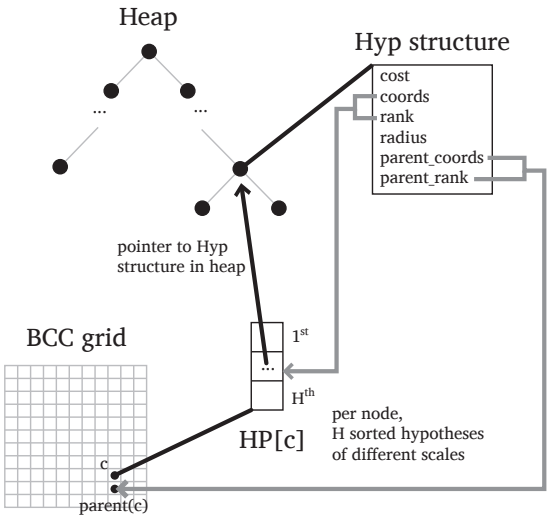


Figure 7.13: Algorithmic structures involved in the implementation of our minimal path algorithm.

7.5 Algorithmic Description and Implementation

7.5.1 Data Structures

Fig. 7.13 illustrates the principal structures involved in the implementation of our algorithm.

`Hyp` structure Hypotheses $x^i_{1:t}$ are represented recursively by `Hyp` structures, encoding:

- *cost*: the cumulative cost $C^*(x^i_{1:t})$ (Eq. 7.9);
- *coords*: the spatial coordinates of the current control point c^i_t;
- *rank*: the current rank of $x^i_{1:t}$ at location c^i_t;
- *radius*: the current radius r^i_t;
- *parent_coords*: the spatial coordinates of c^i_{t-1};
- *parent_rank*: the rank of $x^i_{1:t-1}$, parent of $x^i_{1:t}$, when reaching c^i_{t-1}.

Besides the total cumulative cost $C^*(x^i_{1:t})$, a `Hyp` structure embeds information relative to the current state x^i_t, namely the coordinates of its control point c^i_t and its radius r^i_t. It also stores the current rank of $x^i_{1:t}$ at control position c^i_t. The last two elements, *parent_coords* and *parent_rank*, encode parental links from x^i_t to x^i_{t-1}.

Array `HP` `HP` is a bi-dimensional array maintaining, for each control point location c, a sorted list of the H best hypotheses passing through c. An additional property, motivated in the previous section, is that these H hypotheses all have different scales. The element `HP`$[c][i]$ corresponds to an index pointing to the current $(i+1)^t h$ hypothesis through c (`HP`$[c][0]$ is the current best hypothesis through c). More precisely, `HP`$[c][i]$ is the index of the corresponding `Hyp` structure stored in the `Heap` structure described below.

```
PROPAGATE-PATHS(seed)
    ▷ Initialization
 1  INIT(seed, Heap, HP)
 2  while Heap ≠ ∅
 3        do
              ▷ Extract and pop the head of the heap
 4            hyp_1 ← Heap[0]
 5            POP-HEAP()
              ▷ Explore the neighbors
 6            for each neighbor hyp_2 of hyp_1
 7                 do
                      ▷ Compute cost (Eq. 7.9)
 8                    update_cost ← COMPUTE-UPDATE-COST(hyp_1, hyp_2)
 9                    hyp_2.cost ← hyp_1.cost + update_cost
                      ▷ Attempt insertion on the heap and in HP
10                    INSERT(hyp_2, HP, Heap)
```

Figure 7.14: Pseudo-code of the propagation procedure: PROPAGATE-PATHS.

Heap Heap designates the min-heap structure used to order the different hypotheses globally. It orders Hyp structures in function of their cumulative *cost*, so that its first element is always the current best active hypothesis (of lowest cost). The Heap structure is implemented as an array-based binary heap and is thus addressed in an array-like fashion. In particular, Heap[0] corresponds to the current first element of the heap.

Our approach yields a slight memory overhead compared to textbook implementations of Dijkstra's algorithm. Redundant information is stored in Hyp structures to allow fast, direct access to HP from the heap and conversely (see Fig. 7.13). The current rank and the parental information could potentially be recovered indirectly, at the cost of decreased computational performance. A general remark is that the implementation of our approach is slightly more intricate than classical minimal path algorithms as it has to maintain the set of H hypotheses dynamically, for each considered location. As detailed in Sec. 7.5.4, the memory overhead induced by rich Hyp structures is not critical, due to their temporary presence on the heap and to the possibility of efficiently encoding them.

7.5.2 Algorithmic Procedures

Propagation The pseudo-code for the generic propagation procedure is given in Fig. 7.14. It follows the general design of Dijkstra's algorithm. The variable *seed* typically designates a user-given or automatically detected location at the root of the arterial tree.

The INIT procedure initializes HP and the Heap given the *seed* location. The array HP is initialized with invalid indices (typically of value -1), indicating that all slots $\texttt{HP}[c][h], \forall c, h$ are empty. We consider as initial candidate control points the set $\{c^i_{-1}, /|c^i_{-1} - seed| < D\}$ of grid points close to the *seed* location[3]. For each candidate c^i_{-1}, we derive the set of

3. We use a distance threshold $D = 0.9mm$ in our implementation, allowing some tolerance with respect

```
INSERT(hyp, HP, Heap)
    ▷ Start at the bottom of HP[c]
 1  i ← H − 1
 2  while i ≥ 0
 3        do
 4           if not EMPTY(HP[c][i]) and Heap[HP[c][i]].radius = hyp.radius
 5              then
                      ▷ Same radius hypothesis found
 6                    if hyp.cost < Heap[HP[c][i]].cost
 7                       then
 8                            REPLACE-SAME-RADIUS(hyp, HP, Heap, i)
 9                            return
10                       else
11                            return ▷ hyp is rejected
12           i ← i − 1
    ▷ No same radius hypothesis. Try insertion based on cost
13  INSERT-NEW-RADIUS(hyp, HP, Heap)
```

Figure 7.15: Pseudo-code of the Insertion procedure: INSERT.

initial hypotheses $\{x^{i,j} = (c^i_{-1}, c^{i,j}_0, r^{i,j}_0)\}$ as the product set of the spatial neighbors $c^{i,j}_0$ of c^i_{-1} and all $|R|$ possible radiuses. The initial cumulative cost of each hypothesis is given by:

$$C^*(x^{i,j}_0) = -\log(p^*(r^{i,j}_0)) - \log(p^*_v(y^{i,j}_0|r^{i,j}_0)) + \log(p^*_{bg}(y^{i,j}_0)) + M$$

In the absence of a preferred initial radius, we set $p^*(r^{i,j}_0) = 1$. Each hypothesis is considered for insertion on the heap (procedure INSERT, Fig. 7.15) to initialize the propagation.

The main loop of the algorithm consists in iteratively extracting the first element $hyp_1(c_1, r_1, cost_1)$ from the `Heap`. $hyp_1(c_1, r_1, cost_1)$ is a variable of type `Hyp`, with control point c_1, radius r_1 and cumulative cost $cost_1$. It corresponds to the current best hypothesis on the heap, stored at location `Heap`[0]. Note that in our context, this hypothesis might be the i^{th} hypothesis (with $i \leq H$) for control location c_1, implying that the $i-1$ best hypotheses for c_1 have already been extracted.

In order to augment hyp_1, its neighbors are explored following the neighborhoods described in the previous section. Each neighbor $hyp_2(c_2, r_2)$ corresponds to a new state (p_2, d_2, r_2), with $p_2 = \frac{c_2+c_1}{2}$ and $d_2 = \frac{c_2-c_1}{\|c_2-c_1\|}$ that can be adjoined to hyp_1. The update cost of adding that state to hyp_1 is given by Eq. 7.9. The cumulative cost of the new hypothesis $hyp_1 + hyp_2$ is obtained by adding the update cost to the cumulative cost of hyp_1. The new hypothesis is then checked for insertion by calling the INSERT procedure (line 10).

Selection and Insertion of new hypotheses Once the cumulative cost of a new hypothesis hyp_2 has been computed, the INSERT procedure, given in Fig. 7.15, is used to check whether hyp_2 should be propagated. The INSERT procedure plays a role similar to the RELAXATION procedure in Dijkstra's algorithm (Cormen et al., 2001). It compares the

to the seed placement.

```
REPLACE-SAME-RADIUS(hyp, HP, Heap, cur_rank)
    ▷ Replace hypothesis of same radius at position cur_rank
1   Heap[HP[c][cur_rank]] ← hyp
2   i ← cur_rank − 1
3   while i ≥ 0 and hyp.cost < Heap[HP[c][i]].cost
4         do
               ▷ Update position in HP[c]
5              SWITCH-HP(c, i − 1, i, HP)
6              i ← i − 1
7   i ← i + 1
    ▷ i holds the new rank in HP[c]
8   UPDATE-HEAP(Heap, HP[c][i])
```

Figure 7.16: Pseudo-code of the update/replacement procedure: REPLACE-SAME-RADIUS.

```
INSERT-NEW-RADIUS(hyp, HP, Heap)
     ▷ Start at the bottom of HP[c]
 1   if not EMPTY(HP[c][H − 1]) and hyp.cost ≥ Heap[HP[c][i]].cost
 2      then return ▷ hyp is rejected
     ▷ Replace the worst hypothesis
 3   Heap[HP[c][H − 1]] ← hyp
 4   i ← H − 1
 5   while i ≥ 0 and (EMPTY(HP[c][i]) or hyp.cost < Heap[HP[c][i]].cost)
 6         do
               ▷ Update position in HP
 7             SWITCH-HP(c, i − 1, i, HP)
 8             i ← i − 1
 9   i ← i + 1
     ▷ i holds the new rank in HP
10   UPDATE-HEAP(Heap, HP[c][i])
```

Figure 7.17: Pseudo-code of the insertion procedure: INSERT-NEW-RADIUS.

candidate hypothesis $hyp(c, r, cost)$ with the current H hypotheses recorded for position c, using the HP array. The procedure EMPTY(HP$[c][h]$) checks if the index in HP$[c][h]$ is valid (HP$[c][h] \geq 0$).

INSERT(hyp, HP, Heap) first checks for existing hypotheses of same radius as hyp. If such an hypothesis hyp_2 exists and $hyp.cost \geq hyp_2.cost$ (line 11), hyp is directly rejected as it is not better than hyp_2. If hyp has a lower cost than hyp_2 (line 8), hyp_2 is replaced by hyp. This operation is performed in procedure REPLACE-SAME-RADIUS, given in Fig. 7.16. It maintains the array HP sorted and updates the heap accordingly.

If no hypothesis of same radius as hyp was found (line 13 in Fig. 7.15), hyp can also be incorporated based solely on its cost, by replacing a previous hypothesis of greater cost (procedure INSERT-NEW-RADIUS in Fig. 7.17). The new hypothesis hyp is compared to

the previously (H^{th}) hypothesis stored in $\texttt{HP}[c][H-1]$. If *hyp*'s cost is higher, *hyp* can be directly rejected (line 2). If not, *hyp* can safely overwrite that previously worst hypothesis, which is pruned (line 3). To account for the new hypothesis, the ordering of the hypotheses in $\texttt{HP}[c]$ is corrected and the heap is updated, similarly to Replace-Same-Radius.

In both Replace-Same-Radius and Insert-New-Radius, the procedure Switch-$\texttt{HP}(c, i-1, i, \texttt{HP})$ simply switches the indices in $\texttt{HP}[c][i-1]$ and $\texttt{HP}[c][i]$, following a bubble-sort design. The procedure Update-Heap classically updates the heap structure with respect to the new cost of $\texttt{Heap}[\texttt{HP}[c][i]]$.

Globally, Insert handles the insertion of new hypotheses and maintains both the heap and the HP array. In particular, it ensures that $\texttt{HP}[c]$ is always a sorted list of hypotheses of different scales.

7.5.3 Algorithmic Complexity

The computational complexity of our method is directly derived from that of Dijkstra's original algorithm. For a graph with N vertices, the complexity of Dijkstra's algorithm is $O(N\log(N))$, assuming a binary heap implementation. Our algorithm demultiplies the graph size by keeping H hypotheses per control location. As $H \ll N$, we can neglect the sorting operations involved in maintaining the array HP. The final complexity of the algorithm is thus $O(NH\log(NH))$, where N corresponds in our case to the number of nodes of the 3D spatial grid. One should note that this complexity is of the same order as the existing multiscale minimal path techniques (Wink et al., 2004; Li and Yezzi, 2007; Benmansour et al., 2009) in $O(NS\log(NS))$ with S the number of scales considered. The main difference is that we restrict ourselves to a dynamically adapted subset of S, yielding a lower computational complexity and lower memory requirements (in $O(NS)$ for (Wink et al., 2004; Li and Yezzi, 2007; Benmansour et al., 2009), $O(NH)$ in our case). This makes our approach practically usable on today's workstations.

7.5.4 Implementation Details

Our implementation diverges slightly from textbook descriptions of Dijkstra's algorithm (Cormen et al., 2001). Given the size of our data (the size of the isotropic spatial grid is typically of the order of 512^3 nodes), our particular focus was on limiting memory consumption, so that our algorithm could run on full resolution cardiac CTA data on commercial workstations with 2GB memory limits.

Array HP The role of the array HP is twofold: it stores sorted hypotheses for each control position, enabling hypothesis selection, and provides us with easy access from positions to heap locations. Thanks to HP, the heap can be maintained optimally in terms of memory consumption. In particular, one can note that procedures Replace-Same-Radius (Fig. 7.16) and Insert-New-Radius (Fig. 7.17) systematically reuse the space left by hypotheses that can be safely pruned. Replace-Same-Radius replaces the previous hypothesis of same radius but greater cost, while Insert-New-Radius replaces the H^{th} hypothesis at position c. This way, all the hypotheses on the heap are always valid, active ones. Besides memory considerations, limiting the heap's size has computational benefits, as heap management operations are relatively costly and directly dependent on its size.

The array HP theoretically covers the entire spatial grid, thus having a size of $N \times H$ integers indices. On large grids, the memory space required to store HP becomes prohibitive

for modern commercial workstations. An array of 32 bit integers corresponding to a 512^3 grid with $H = 4$ would occupy 2GB of memory. Fortunately, propagation is usually limited to a very small portion (up to a few percents) of the grid, as demonstrated in Sec. 7.6.1 and 7.6.2. We exploit this property by allocating the `HP` array only for the portions of space actually visited. We use a block volume structure (see Sec. 5.3 and Fig. 5.4), with blocks of 16^3 locations.

Storing information on the heap As was illustrated in Fig. 7.2, the propagation scheme can be viewed as a process evolving a *front* interface across the graph. The active front actually corresponds to the hypotheses on the heap, whose accumulative cost is being optimized. As an hyper-surface, the front is thus of much smaller size than the entire image. It is also generally smaller than the set of explored, optimized positions in `HP`. Our approach to minimizing memory consumption is thus to store most of the required information in heap elements, the `Hyp` structures detailed in Fig. 7.13.

The naive implementation of `Hyp` structures can be relatively costly in memory. Fortunately, their storage can be greatly optimized. As our spatial grids do not exceed the size of $512 \times 512 \times 1024$ positions, *coords* can be encoded using a 28 bit lexicographic index. The *radius* information can be encoded through its index in the discrete set R. In our implementation, we use 13 different radiuses. A radius index thus requires only 4 bits. Similarly, the parent coordinates (*parent_coords*) are encoded as the index of the corresponding 14-connected BCC neighbor, yielding a cost of 4 additional bits. Finally, the *rank* and *parent_rank* indices are always inferior to H, which does not exceeds 4 (hence stored on 2 bits) in our implementation. Besides the cumulative cost (a 32 bit floating point value), we theoretically only need another $28 + 4 + 4 + 2 + 2 = 40$ bits to store all the required information. In our implementation, we use a 32 bit integer to encode both *coords* and *radiuses*, and an additional 8-bit integer for current rank parental information, yielding virtually no storage overhead per `Hyp` structure. Information clustered in the different integers is simply retrieved through bit shifting and bit masking operations.

Re-using `HP` to store path retrieval information As an hypothesis hyp_1 is retrieved from the heap (line 4 in the PROPAGATE-PATHS procedure in Fig. 7.14), its cumulative cost is definitive. In classical front propagation techniques, a popped node enters the set of optimized nodes (see Fig. 7.2). In our context, if hyp_1 corresponds to the i^{th} hypothesis for position c, it means that no unseen hypothesis arriving in c in the future can outrank it. As hyp_1 exits the heap, the heap index in $\texttt{HP}[c][i]$ is not useful anymore. Before actually popping hyp_1, we encode the information of radius, parental coordinates and rank in $\texttt{HP}[c][i]$. This data totals only 10 bits. This way, hyp_1 can be safely deleted from the heap, freeing up memory, and the space in $\texttt{HP}[c][i]$ is re-used. At convergence, the array `HP` encodes local radius and parental links, allowing the full retrieval of paths from any control position and rank. One can note that the *cost* information is not required and is actually not available anymore. This approach differs from classical implementations which maintain the action (cost) map permanently. In our implementation, the only permanent structure is the array `HP`.

Possible algorithmic optimizations Our current implementation could potentially be optimized in a number of ways. First, one could try to limit redundant computations. A particular state $x_t^i = (p_t^i, d_t^i, r_t^i) \equiv (c_t^i, c_{t-1}^i, r_t^i)$ can be reached from multiple predecessors. This implies that the corresponding observation y_t^i, used to value the likelihood terms in

Eq. 7.9, can be queried multiple times, yielding multiple `MFlux` computations with the same parameters. One could imagine a cache structure to avoid such redundant computation. In our experiments, we were unfortunately unsuccessful in attaining a satisfying balance between computational gain and additional memory requirements.

A second lead for computational improvement could be the use of a more efficient priority queue, for instance switching from a binary heap to a bucket-based priority queue, following the ideas from (Yatziv et al., 2006; Falcao et al., 2000). This would however require an adaptation of the `HP` pointer array.

7.6 Workflow Description

7.6.1 Interactive Workflow

The most direct application of our algorithm is a classical two-point interactive framework, where seed and end locations (noted respectively c_0 and c_{end}) are user-provided. In our coronary artery application, we assume the seed to be placed at a more proximal location than the end point. In particular, full branch extraction is performed by placing the seed at the ostium (origin of the coronary tree from the aorta) and the end point at the distal end of the target branch.

As previously mentioned, the seed is used to initialize the propagation procedure and serves as the seminal control point location c_0, from which hypotheses in all directions and all scales are created and propagated. A minor modification to the general propagation procedure (Fig. 7.14) is to stop evaluations as soon as an hypothesis with location c_{end} is extracted from the heap (line 4). Indeed, the first hypothesis extracted for c_{end}, irrespectively of its radius, is the best possible path from the seed to c_{end}. No new hypothesis considered afterwards can have a lower cumulative cost. This way, only a fraction of the search space is actually explored, as will be illustrated in Sec. 7.7. This approach is usually referred to as *partial front propagation* in the literature (Deschamps, 2001).

Finally, notice that this workflow is not necessarily limited to two-point extractions. One can easily specify several end points for a single seed and all branches will be extracted simultaneously. In this scenario, the propagation stops when hypotheses for all end points have been found.

7.6.2 Full Tree Automatic Extraction

We now describe how we can supplement and adapt our algorithm to obtain a fully automatic extraction of both left and right coronary trees. The idea is to automatically detect ostia locations to initialize the minimal path optimization process. This process is now used as a propagation scheme, for which a stopping criterion is derived. Finally, the tree structure is extracted retrospectively from the propagated paths.

7.6.2.1 Automatic Aorta Segmentation and Ostia Detection

To automatize the detection of the seed points, *i.e.*, the ostia points, we rely on existing tools developed at Siemens Corporate Research. Automatic aorta segmentation is obtained through a fast variant (Grady, 2006) of the isoperimetric graph-based algorithm (Grady and Schwartz, 2006). Ostia locations are detected through a simple growing process illustrated in Fig. 7.18. We start our propagation algorithm from the surface of the segmented aorta and retain the propagated hypotheses first reaching a certain distance

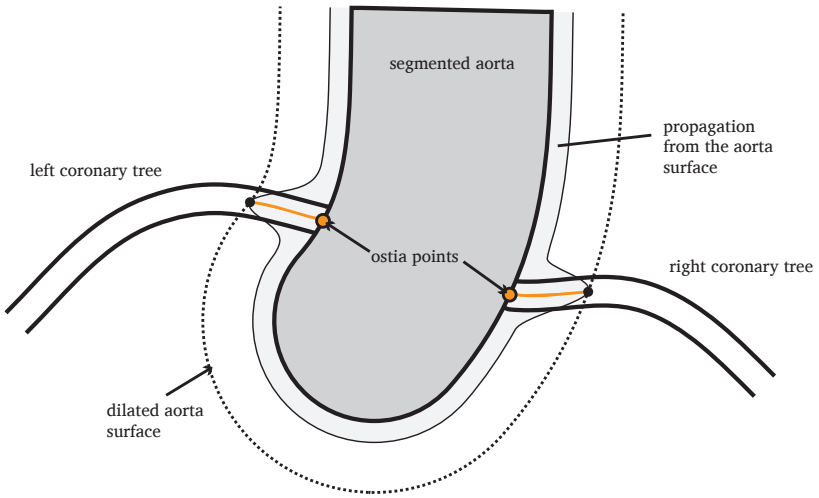


Figure 7.18: Automatic ostia detection. The propagation process is started from the automatically segmented surface of aorta. The first hypotheses reaching a distance threshold from the aorta are backtracked to obtain the ostia points.

(typically 1cm) from the aorta's surface. Hypotheses will propagate faster in the coronaries, reaching the distance threshold faster. We keep track of the two main connected components of the propagation front reaching the critical distance, one per arterial tree. The first (thus best) hypothesis per connected component is backtracked up to the aorta's surface to determine the corresponding ostia point.

This scheme already proved its excellent reliability in previously published works (Dikici et al., 2008; Tek et al., 2008; Gulsun and Tek, 2008b). A typical result is illustrated in Fig. 7.19. Notice that we could potentially continue the propagation from the whole aorta surface and extract both trees simultaneously. The definition of the ostia points allows us to separate the trees more easily and potentially parallelize the execution.

Propagation is started from each detected ostium separately. The aorta mask is reused during the propagation itself, to prevent hypotheses from propagating back to the aorta. Additionally, we use the graph-cut-based automatic heart isolation algorithm from (Funka-Lea et al., 2006) to constrain the exploration within the heart and exclude, in particular, lung areas (see Fig. 7.20). This mask reduces the overall search space and effectively prevents the algorithm from extracting pulmonary arteries. In practice, each newly considered hypothesis hyp_2 is checked before evaluation (line 8 in Fig. 7.14). If hyp_2 occupies a location within the mask, *i.e.*, within the aorta or outside the heart, it is rejected. Additionally, the mask is updated during the propagation to include locations where all H hypotheses have already been popped. It is indeed useless to revisit such locations, as no new hypothesis could be accepted. Re-using the mask provides us with an elegant way of avoiding redundant evaluations.

7.6.2.2 Stopping Criterion

In the absence of a user-provided end point, defining an adequate stopping criterion is a key, challenging task. We propose to exploit the probabilistic semantics of our cost metric.

Figure 7.19: Sample result of automatic aorta segmentation and ostia detection. On this 3D maximal intensity projection view of a cardiac CTA dataset, the segmented aorta is highlighted, with the two (left and right) detected ostia points plotted as red squares.

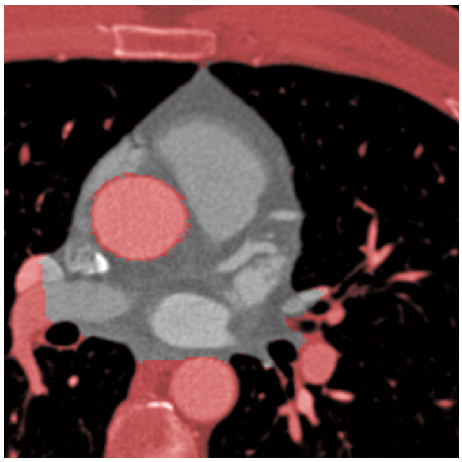


Figure 7.20: Fusion of the automatic aorta and heart segmentations to constrain the search space. Excluded areas (aorta and positions outside the heart) are highlighted in red on this MPR view of a cardiac CTA dataset.

We implemented an approach based on a threshold on the update cost (Eq. 7.9). The idea is to prevent hypotheses from propagating in non-vascular areas, *i.e.*, parts of the search space of low probability. We propose to use a threshold T on the cost $C(x^i_{t-1}, x^i_t) = -\log(p(r^i_t|r^i_{t-1})) - \log(p_v(y^i_t|r^i_t)) + \log(p_{bg}(y^i_t)) + M$ over which an hypothesis is considered as most likely to be propagating in non-vascular regions. This threshold T is defined as:

$$T = T_l + T_p + M$$

The term T_l corresponds to a constraint on likelihood responses. Intuitively, we want

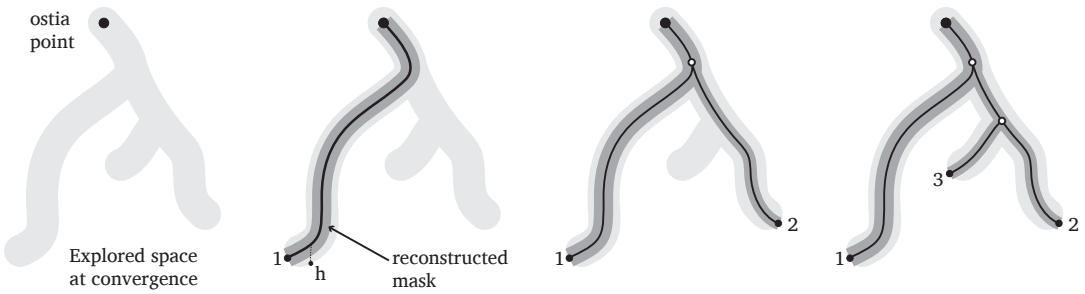


Figure 7.21: Hierarchical tree extraction. From left to right: from the explored space at convergence, branches are recovered iteratively. See text for details.

the image observations to be more likely to come from vessels than from the background. We set T_l to correspond to $p_v(y_t^i|r_t^i) = p_{bg}(y_t^i)$, which simply gives $-\log(p_v(y_t^i|r_t^i)) + \log(p_{bg}(y_t^i)) = 0$. As illustrated in Fig. 6.4, this would correspond to a direct threshold on the `MFlux` feature response (where the two distributions intersect). Notice that the corresponding feature threshold is both radius- and dataset-dependent, because of $p_v(y_t^i|r_t^i)$ and $p_{bg}(y_t^i)$, respectively. The term T_l is set to discriminate hypotheses gathering low likelihood responses. Conjointly, T_p prevents hypotheses with too low priors $p(r_t^i|r_{t-1}^i)$ from surviving. We empirically set $T_p = -\log(0.15)$ (see Fig. 7.11). Our tests showed that the final result is not overly sensitive to changes of T_p. Our threshold on the update cost is finally defined as:

$$T = T_l + T_p + M = -\log(0.15) + M$$

Setting a strict threshold on a single-edge update cost is likely to cause premature stopping of the propagation because of local image perturbations (noise, artifacts or pathological anomalies). For our stopping criterion to be more robust, we base our analysis on the S last states/edges $[\![x_{1:t-S}^i, x_{1:t}^i]\!]$ of an hypothesis $x_{1:t}^i$. If more than half the update costs for the last S edges fall over T, the hypothesis is halted. We used $S = 20$ in our experiments, corresponding to an approximate centerline length of $6mm$ with a grid resolution of 0.3mm. This value proved to be a good tradeoff between the prevention of premature stopping and leaking issues (false positives).

For each of the automatically detected ostia locations, the general propagation procedure (Fig. 7.14) is performed until no hypothesis passes the stopping criterion. In the following section, we detail how the tree structure is extracted from the set of propagated hypotheses.

7.6.2.3 Hierarchical Tree Extraction

At convergence, the tree structure is extracted in a hierarchical fashion similar to the method of (Gulsun and Tek, 2008b). Fig. 7.21 illustrates this process. Halted hypotheses lie at the frontier of the explored space. According to our criterion, hypotheses were stopped because they were unlikely to correspond an actual vessel anymore. This can have two causes: hypotheses either diverged from an existing artery, or reached its distal end. Intuitively, hypotheses reaching the end of an artery are the longest ones for that particular artery.

Halted hypotheses are ordered with respect to their geodesic distance from the ostia. Approximating this distance by the length L of the hypothesis $x^i_{1:L}$ proved to be sufficient in our experiments. The hypothesis of longest length (hypothesis 1 in Fig. 7.21) is assumed to correspond to the longest branch and serves as the first element of the reconstructed tree. We exclude the S last states, to account for the tolerance of the stopping criterion (see previous section). Thanks to the radiuses r^i_t associated to the hypothesis, we construct a volumic mask as the union of discrete spheres of center p^i_t and radius r^i_t. This mask approximates the spatial coverage of the hypothesis.

The process then evaluates hypotheses in decreasing order of length. Each hypothesis is considered for inclusion in the tree structure. An hypothesis h is rejected if its centerline length *outside* the current mask is lower than $2S$. This scheme effectively prunes the numerous spurious "sideways" hypotheses which geometrically do not correspond to a significant new branch, *i.e.* hypotheses which were stopped as they exited the vessel lumen. When a new branch is discovered (hypothesis 2 in Fig. 7.21), it is added to the tree structure. A branching point is assessed as the location where the backtracked path from hypothesis 2 arrives sufficiently close to a previous centerline in the tree. We use a threshold of 0.6mm, about 2 voxels, in our implementation. The mask is updated with each newly accepted hypothesis, and the process continues until all hypotheses have been considered.

This simple and efficient hierarchical scheme copes elegantly with bifurcations by detecting them retrospectively. In contrast, direct tracking techniques (see Sec. 3.3.4 and 3.3.5 and Chapter 8) often try to detect branchings as they occur, which can be relatively delicate and/or involve additional procedures.

We eventually obtain a tree structure composed of a low number of hypotheses (up to 20 in practice) corresponding to the identified coronary branches. This result representation is particularly compact and informative, encoding both centerline curves and associated radiuses. Compared to the extraction scheme from (Gulsun and Tek, 2008b), our principal contribution is the use of the volume information associated to the hypotheses to derive an intuitive pruning criterion based on their overlap with the partially reconstructed tree.

7.7 Evaluation

7.7.1 Qualitative Evaluation

Robustness Our algorithm demonstrated its robustness in various conditions of acquisition quality, noise and image artifacts, and in presence of pathologies such as calcifications (Fig. 7.22), stenoses and stents (Fig. 7.23). The two-point interactive workflow proved its adequacy for the segmentation of chronic total occlusions (Fig. 7.24). In such cases, the image content (contrast in particular) is generally not sufficient to guide the segmentation process. The model-based information (radius prior in particular) plays a key role in the robust extraction of such pathologies. On the other hand, the stopping criterion used for automatic tree extraction can lead to premature stopping in most severe cases, depending on the parameter S. The ability of the automatic process to bridge severe occlusions is to be balanced with potential leakage issues.

Our space+scale approach also demonstrated its capacity to extract coronaries of various radiuses and curvature, from large, relatively straight segments to curvy, thin distal branches (see Fig. 7.25). Fig. 7.26 and 7.27 show samples of coronary trees extracted fully automatically, to be compared with the ground-truth manual delineation. In a vast

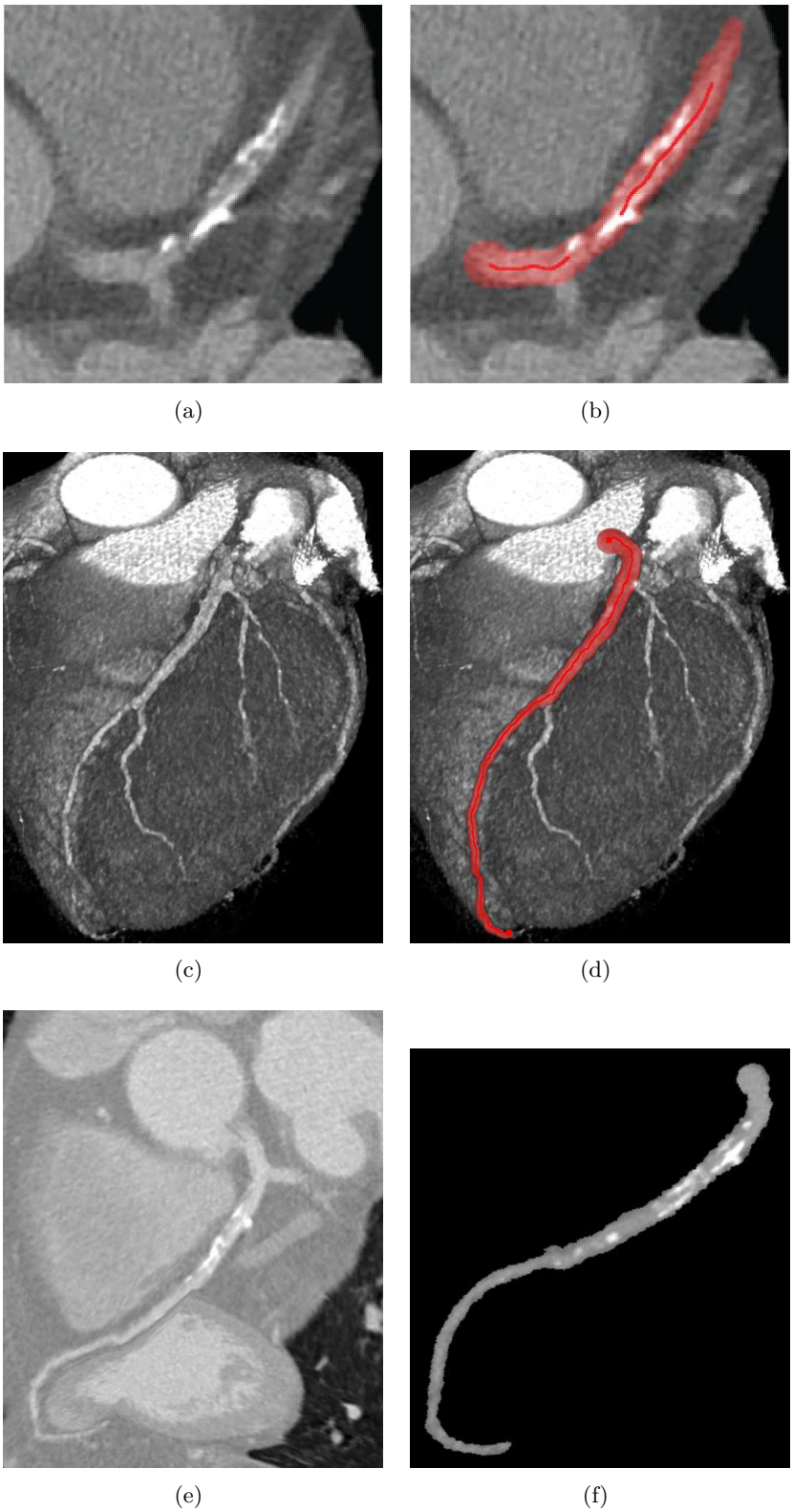


Figure 7.22: Extraction of a severely calcified right coronary artery (two-point workflow, $H = 4$). Volume rendering (a) and MPR view (c) of the original data. (b) and (d) Extracted vessel overlayed. (e) CPR view of the extracted branch. (f) Local tube MIP view of the extracted branch.

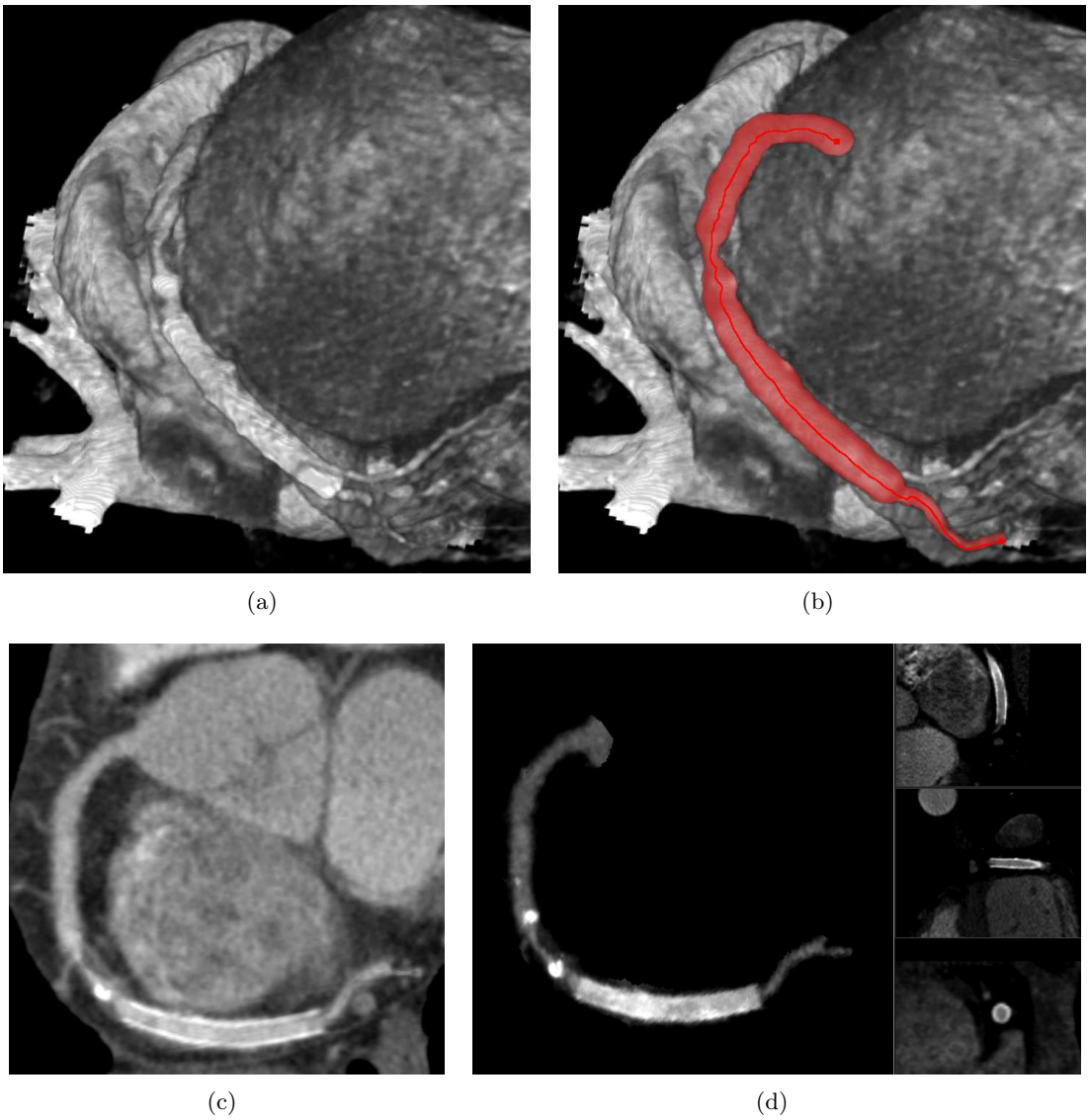


Figure 7.23: Extraction of a right coronary artery with a stent, calcified plaque and stenosis. (two-point workflow, $H = 4$). (a) Volume rendering of the original data (right coronary tree). (b) Extracted vessel. (c) CPR view of the extracted branch. (d) Local tube MIP view of the extracted branch, with orthogonal MPR views of the stent.

majority of cases, the process successfully extracts all main branches and most secondary, thinner ones. The algorithm may under-segment thin distal parts of the coronaries, particularly in noisy datasets. Such vascular segments are arguably of lower medical relevance. Conversely, the automatic algorithm was sometimes able to extract vessels deeper than the human experts from visual inspection, or even to discover branches not tagged in the manual delineation.

The fully automatic workflow is relatively robust to leakage issues. As illustrated in Fig. 7.28, the sparse leakage occurrences we encountered can be classified into two categories. The first type of leakage was mainly observed on the circumflex branch of

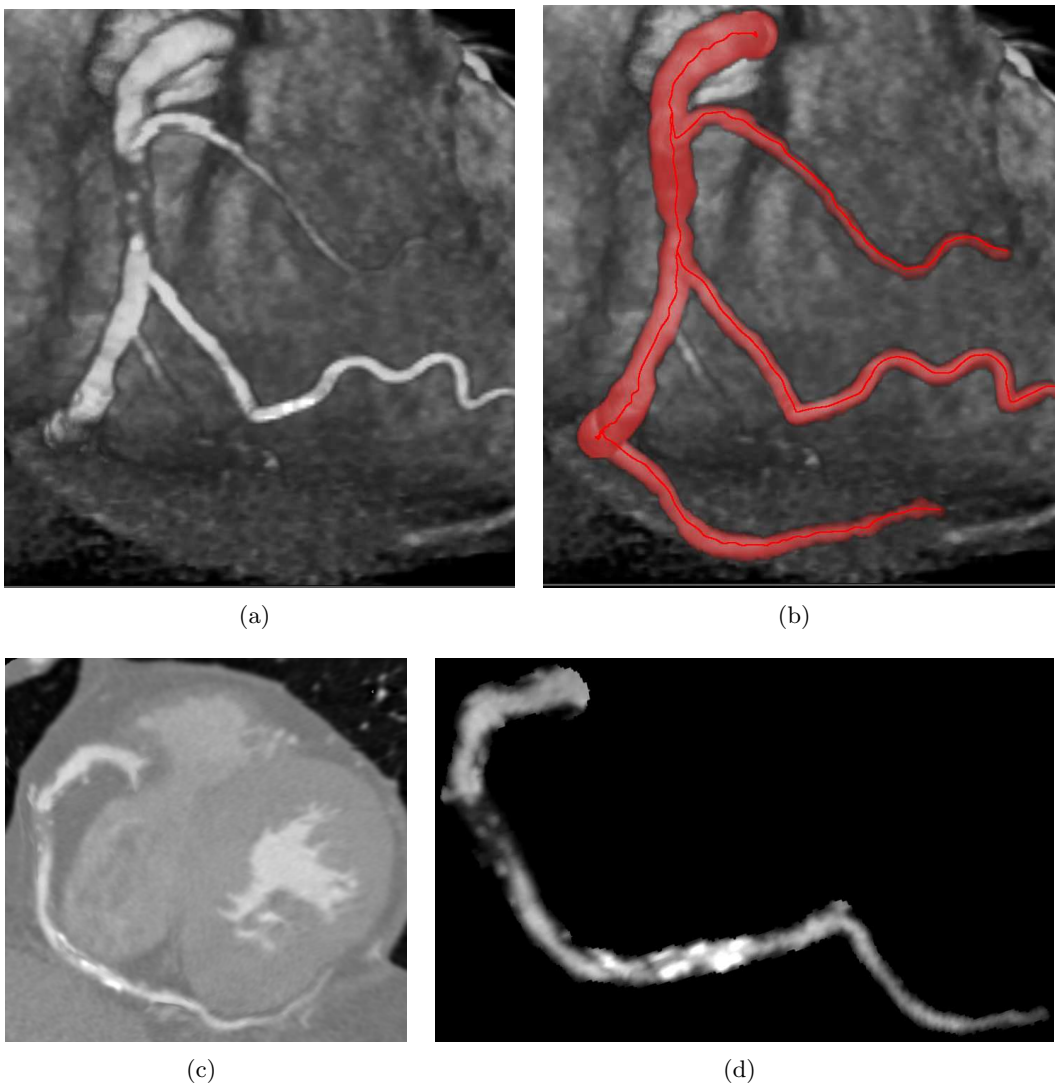


Figure 7.24: Extraction of a left coronary tree with total occlusion (interactive workflow, 1 ostium seed, 3 distal seeds, $H = 4$). (a) Volume rendering of the original data. (b) Extracted vessel. (c) CPR view of the extracted branch. (d) Local tube MIP view of the extracted branch.

left coronary trees, where arteries run close to the ventricle. In such cases, the automatic process sometimes jumped in the ventricle, considering its elongated tip as a potential branch. One can notice that the propagation into the ventricle remains limited, as the estimated radius quickly exceeds that of a potential coronary artery. We might be able to automatically remove such leakage through radius- or volume-based pruning as a post-processing step (see for instance (Gulsun and Tek, 2008a)). The second category, that is leakage into veins, is more problematic. In particular, the distal ends of right coronaries (RCA) may run really closely, often parallely to the vein network. Our algorithm may then jump and propagate hypotheses into the veins. The resolution of such issues is much more intricate, as the spurious branches are actual vessels, with intensity and geometric properties similar to the target coronaries. A more global approach, including anatomical

Figure 7.25: Illustration of scale adaptation. Extraction of a thin artery branching off the main right coronary (two-point workflow, $H = 4$). Left column: volume rendering and MPR view of the original data. Right: extracted vessel.

information about both networks, is likely to be required to resolve such ambiguities. Potential leakage into veins is the counterpart of the robustness of the approach. The tradeoff between potential under- and over-segmentation is controlled by parameters T and S. One could increase T and decrease S to obtain more local, less tolerant stopping criteria. We emphasize that from a clinical point of view, we believe that it is more valuable to limit false negatives, *i.e.*, to provide experts with extraction results as complete as possible. False positives such as extracted veins are easily rejected upon visual inspection by human experts, whereas faint coronary continuations are more difficult to assess by manual delineation.

Finally, shortcut issues were observed mainly on faint, thin and curvy vessels. They remained spatially limited in a vast majority of cases. In this regard, we believe that the careful design of our cost metric is the key factor of the robustness of the method to potential shortcuts. The quantitative evaluation presented in the next section corroborates our qualitative observations.

Computational performance Thanks to its careful algorithmic design, our approach is particularly efficient, making it exploitable on today's commercial workstations. The computational effort involved in extracting a single artery or a full tree depends on several factors. First, it is relative to the extent of required exploration, which varies with the quality of the acquisition and with the extent and complexity of the vascular network. Second, execution times are directly related to the choice of parameter H (algorithmic complexity of $O(NH\log(HN))$). With our C++ implementation of the two-point interactive workflow, the average vessel branch is extracted in about one minute for $H = 4$ (less than 15 sec. for $H = 1$) on a 2.16GHz `Intel Core Duo Intel` processor. The time required to automatically extract a full arterial tree is of the same order.

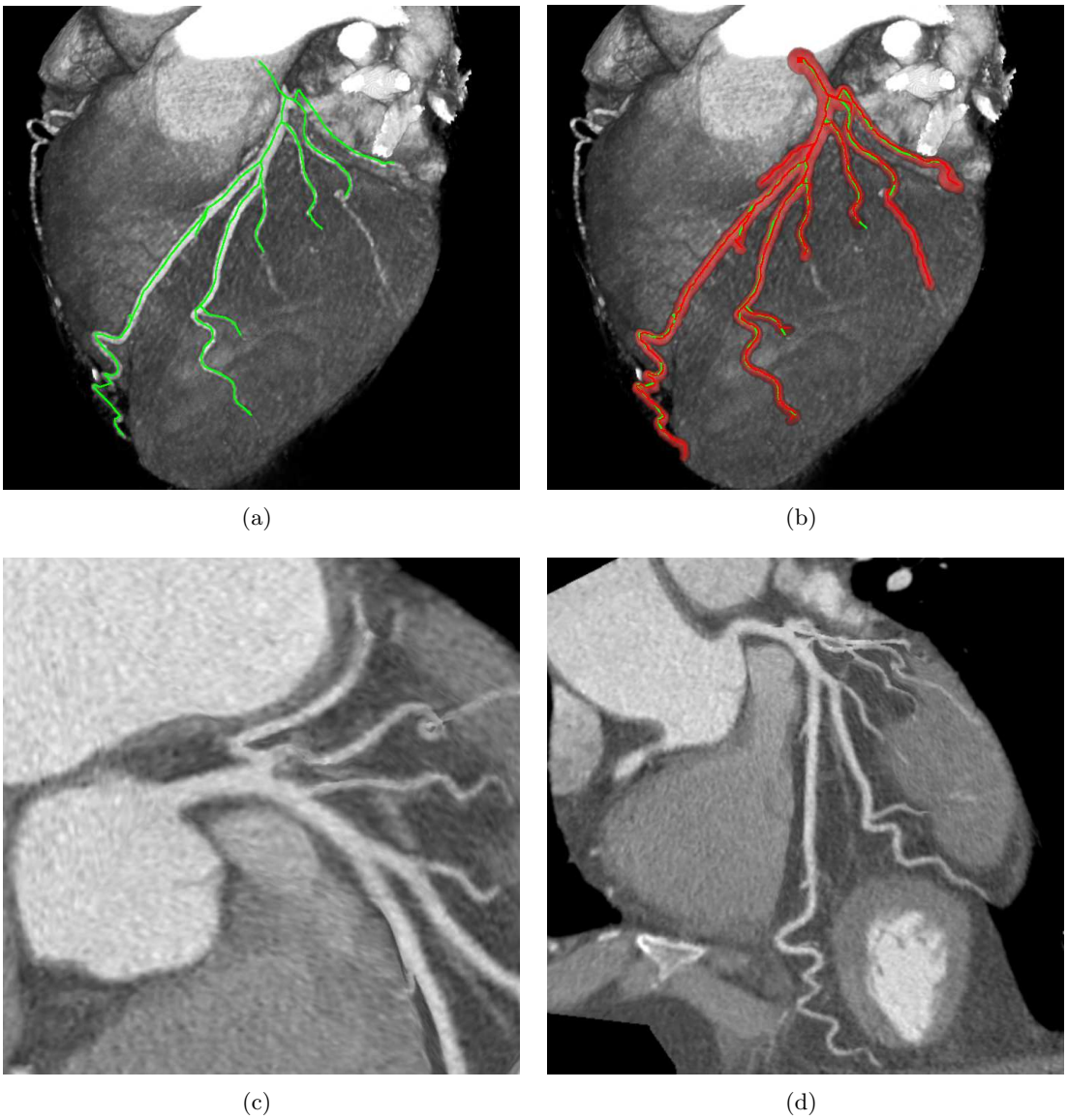


Figure 7.26: Minimal path extraction: sample of full tree, automatic extraction ($H = 4$). (a) Volume rendering of the original data (left coronary tree) with ground-truth centerline in green. (b) Automatically extracted tree overlayed. (c) and (d) Two examples of CPR views generated from the automatic result.

Search space and background likelihood In order to extract the longest arterial branch of a coronary tree, the two-point version of our algorithm will basically explore the entire arterial tree (see Fig. 7.29). Notice that to ensure the extraction of a path between any two points, no stopping criterion is used, so that paths are propagated even in non-vascular areas. The extent of exploration required is directly dependent on the discriminative power of the cost metric, *i.e.*, its ability to favor correct hypotheses over erroneous ones. In this regard, the use of the background likelihood term in Eq. 7.9 has a limited impact on the extracted result, but markedly decreases the extent of exploration

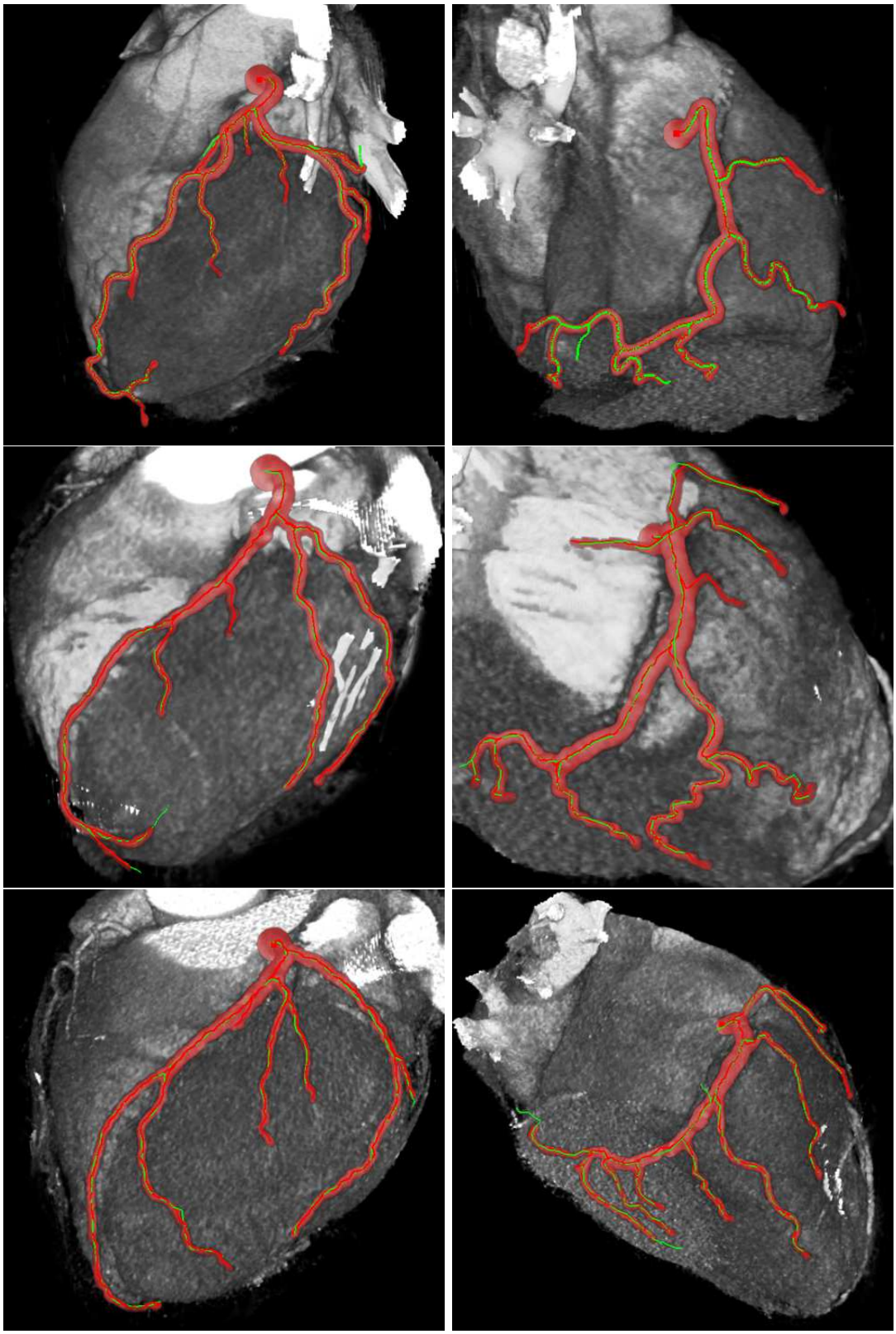


Figure 7.27: Minimal path extraction: additional samples of full tree, automatic extraction ($H = 4$). Left column: left coronary trees. Right: right coronary trees.

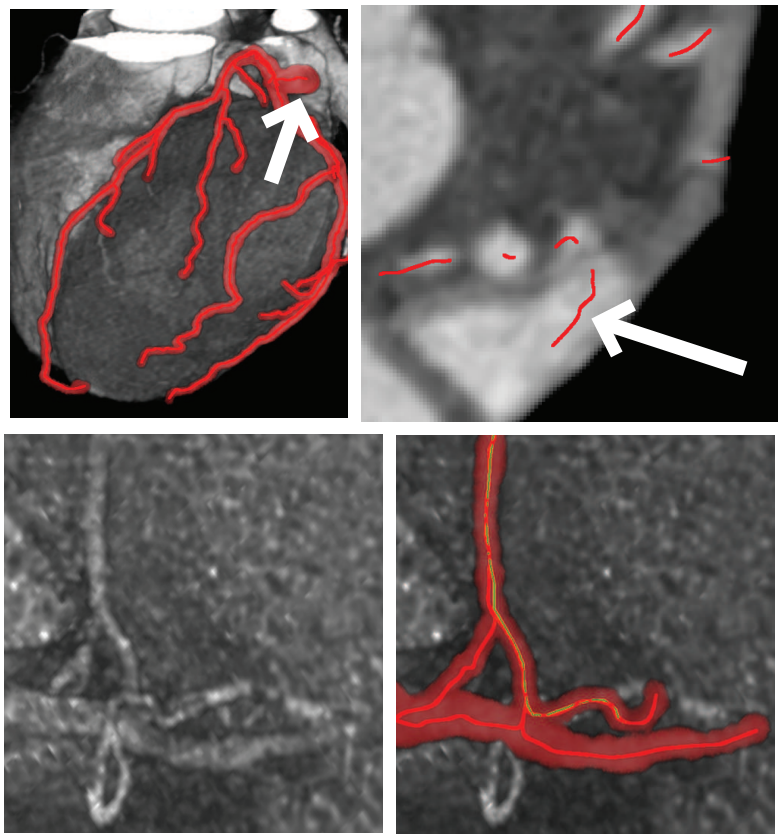


Figure 7.28: Leakage issues. Top row: volume rendering (left) and MPR view (right) of leakage in the elongated tips of the ventricles. Bottom row: leakage in the coronary sinus. Original data is shown on the left, the extracted result on the right. Ground truth delineation is given in green.

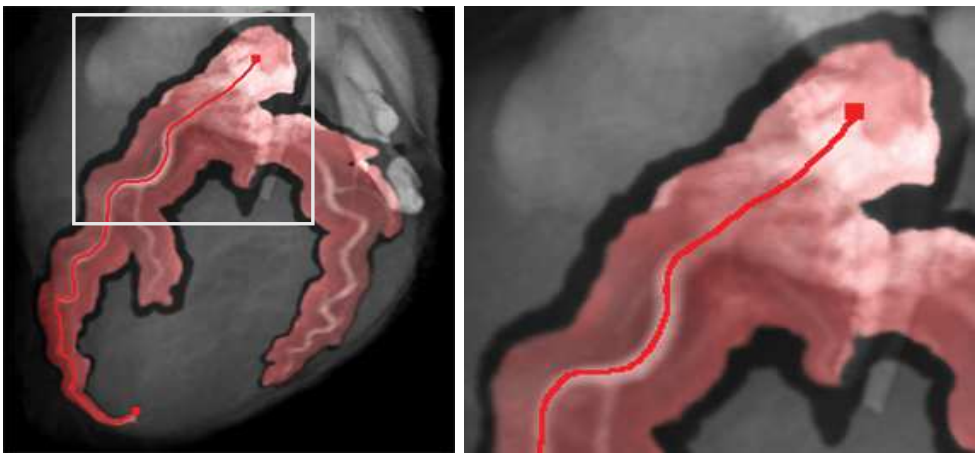


Figure 7.29: Search space and influence of the background likelihood term ($\log(p_{bg}(y_t^i))$ in Eq. 7.9). On the left, we highlight the portion of the search space explored with (red, 1.1×10^6 positions) and without (dark, 1.95×10^6 positions) background likelihood term. Right: detail of the left image.

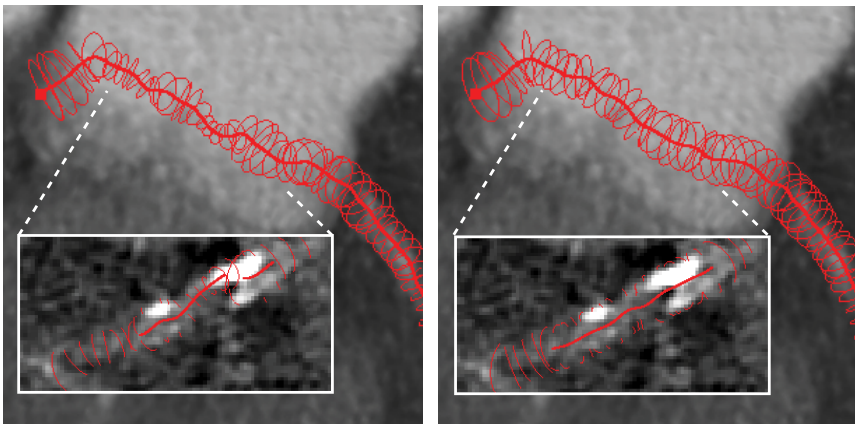


Figure 7.30: Influence of the radius prior term ($-\log(p(r_t^i|r_{t-1}^i))$ in Eq. 7.9) on the robustness of the extraction process. Illustration for a relatively low image quality, heavily calcified dataset. Left: extraction without radius prior. Right: extraction with radius prior.

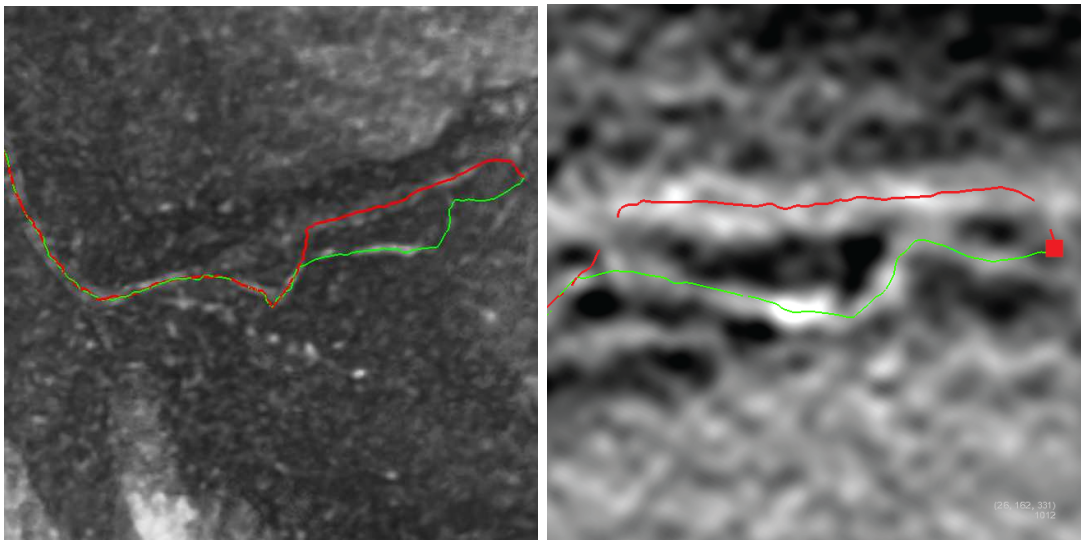


Figure 7.31: Influence of the parameter H. Detail of the extraction of the thin distal end of a right coronary artery. Left: volume rendering. Right: MPR view. In red, the vessel extracted with $H = 1$ jumps into a vein and thus misses the target vessel. For $H = 4$, the correct artery was extracted (green line).

by penalizing hypotheses in non-vascular areas. The gain can reach ∼50% in practice (see Fig. 7.29).

Radius prior $p(r_t^i|r_{t-1}^i)$ The scale prior term $-\log(p(r_t^i|r_{t-1}^i))$ improves the robustness in presence of local anomalies and/or low image quality as illustrated in Fig. 7.30, while still allowing smooth scale adaptations (Fig. 7.25).

Measure	H=1	H=2	H=4
FN	1.6%	1%	0.8%
FP	1.5%	0.8%	0.6%
OV	98.4%	99.1%	99.3%
OT	98.9%	99.2%	99.4%
AI	0.31mm (±0.25)	0.29mm (±0.25)	0.29mm (±0.24)
AR	0.19mm (±0.18)	0.19mm (±0.18)	0.19mm (±0.18)

Table 7.1: Minimal path extraction: two-point interactive segmentation. Internal quantitative validation. See text for details.

Measure	% / mm			score		
	min.	max.	avg.	min.	max.	avg.
OV	74.4%	100.0%	97.6%	38.9	100.0	81.3
OT	74.6%	100.0%	97.7%	38.8	100.0	81.3
AI	0.21 mm	0.54 mm	0.29 mm	26.7	57.0	36.9

Table 7.2: Minimal path extraction: two-point interactive segmentation ($H = 4$). Summary of Rotterdam evaluation (24 testing datasets).

Number H of kept hypotheses The effect of the parameter H is delicate to analyze and illustrate, as it is intimately linked to that of the scale prior. H controls the strictness and, indirectly, the locality of hypothesis pruning. In most cases, the results obtained for different H values are very similar, confirming the assumption that radius selection can generally be performed very locally ($H = 1$) in a robust manner. Differences appear in areas where the radius information is more ambiguous, because of noise, lack of contrast or pathologies. Fig. 7.31 gives an example where using $H = 4$ led to the correct extracted path, whereas $H = 1$ yielded an erroneous, shortcut result.

The greater H, the closer the exploration is to the globally optimal, full scale-space graph exploration. The quantitative evaluation of the next section indicates that increasing H leads to better robustness on average, confirming the adequacy of our cost metric and optimization scheme.

Algorithm	OV	OT	AI
Our approach	97.6%	97.7%	0.29mm
MHT (Friman et al., 2008b)	98.5%	98.7%	0.23mm
Tracer (Szymczak, 2008)	95.1%	95.5%	0.26mm
TwoPointMinCost (Metz et al., 2008a)	91.9%	92.5%	0.46mm

Table 7.3: Minimal path extraction: two-point interactive segmentation ($H = 4$). Quantitative comparison with other methods on the Rotterdam Evaluation Framework (testing database, average results).

FN	FP	OV	OT	AI	AR
13.5%	16.5%	85%	92.9%	0.31mm (±0.25)	0.2mm (±0.18)

Table 7.4: Minimal path extraction: full tree automatic extraction. Internal quantitative validation ($H = 4$). See text for details.

Measure	**% / mm**			**score**		
	min.	**max.**	**avg.**	**min.**	**max.**	**avg.**
OV	58.9%	99.7%	91.3%	29.5	96.7	50.9
OT	58.9%	100.0%	93.9%	29.5	100.0	67.7
AI	0.21 mm	0.56 mm	0.31 mm	24.1	61.5	36.3

Table 7.5: Minimal path extraction: full tree automatic extraction ($H = 4$). Summary of Rotterdam evaluation (24 testing datasets).

7.7.2 Quantitative Evaluation

7.7.2.1 Interactive Workflow

Our interactive implementation was first tested on our ground-truth database (see Appendix B). We used 10 randomly selected datasets for training (learning of the prior and likelihood distributions, see Sec. 6) and the 51 other datasets for testing. For each dataset and for both left and right coronary trees, we performed the extraction of every branch (from the ostium to the distal end) present in the manual delineation. This totals over 800 vessel segments, from main coronary branches to thin secondary branches.

Both ground truth delineations and extracted results were resampled, using cardinal splines, at a resolution of $0.03mm$, to allow point to point comparisons. For our evaluation, we followed the guidelines of (Schaap et al., 2009a). A point of the ground truth (resp. extracted result) is marked as true positive (TP_{gt}, resp. TP_{res}) if it lies within ground truth radius distance of a point of the extracted result (resp. ground truth centerline). If not, the point of the ground truth centerline (resp. extracted result) is marked as false negative (resp. false positive).

Table 7.1 gives the results of our quantitative evaluation. Criteria include:

- FN: average percentage of false negatives, *i.e.*, portion of the ground truth delineation not successfully extracted by the algorithm;
- FP: average percentage of false positives, *i.e.*, portion of the extracted result not

Algorithm	OV	OT	AI
Our approach	91.3%	93.9%	0.31mm
GVFTube'n'Linkage (Bauer and Bischof, 2008a)	92.7%	95.3%	0.37mm
DepthFirstModelFit (Zambal et al., 2008)	84.7%	87%	0.28mm
AutoCoronaryTree (Tek et al., 2008)	84.7%	86.2%	0.34mm

Table 7.6: Minimal path extraction: full tree automatic extraction ($H = 4$). Quantitative comparison with other methods on the Rotterdam Evaluation Framework (testing database, average results).

present in the ground truth delineation;
- OV: average overlap ratio (see below);
- OT: average overlap ratio for clinically relevant vessels (diameter $\geq$ 1.5mm);
- AI: average centerline error (distance to the ground truth, for true positive portions of the result, standard deviation in parentheses);
- AR: average radius estimation error (for true positive portions of the result, standard deviation in parentheses);

The first four measures evaluate the robustness of the method, the last two its centerline and radius accuracy.

The overlap ratio OV is defined as:

$$OV = \frac{TP_{res} + TP_{gt}}{TP_{res} + TP_{gt} + FN + FP}$$

where TP_{res} and TP_{gt} are the percentages of true positives for the extracted result and ground truth delineation, respectively. The ratio OV corresponds to a Dice coefficient measuring the similarity between the result and the ground truth. The ratio OT is defined similarly to OV, but restricted to clinically relevant arteries, of a diameter greater than 1.5mm (Schaap et al., 2009a).

Our algorithm was evaluated for different values of H. For $H = 1$, the algorithm already proves its high robustness (average OV of 98.4%). This experiment shows that increasing H further improves the robustness of the extraction, confirming the relevance of our model and heuristics. For values greater than 4, the improvement observed was unsignificant and outweighed by the additional time and memory requirements. The average centerline distance AI is slightly larger than the typical intra-slice data resolution (0.3mm per voxel). This accuracy result is coherent with the discrete nature of our algorithm. The average radius estimation error AR of 0.191mm is satisfyingly under voxel resolution.

We additionally submitted results (for $H = 4$) to the publicly open Rotterdam Coronary Artery Algorithm Evaluation Framework of the Erasmus Medical Center (Schaap et al., 2009a). The motivation is to make our approach directly comparable to other methods and to benefit from the intra-observer measures included in the evaluation. The evaluation includes 8 training and 24 testing datasets. Table 7.2 summarizes the results obtained on the testing database. Fully detailed results, including on their training database (not used for our training procedure) are given in Appendix D. For each dataset, the Rotterdam evaluation considers four main vessels manually delineated by experts, totaling 96 vessels for the 24 testing datasets. We emphasize that the ground truth delineation for the testing data is not provided. Given submitted extraction results (centerline curves), the evaluation is performed on the servers of the Erasmus Medical Center. Besides the measures we already presented, this evaluation includes scoring with respect to the inter-observer variability. Each score is defined in $[0, 100]$ (Schaap et al., 2009a). A score of 50 corresponds to a result of the order of the inter-observer variability, 100 to a perfect result.

The good performance of our method is confirmed, with 97.6% average overlap, exceeding inter-observer overlap (score of 81.3). The average distance (AI) in Table 7.2 is slightly larger than inter-observer variability (score of $36.9 < 50$). Our primary focus being robustness, we consider this accuracy level to be satisfactory for initial delineation before subsequent refinement. One can notice that the overlap result is slightly lower than the one obtained on our internal evaluation. The Rotterdam evaluation includes numerous datasets of low acquisition quality and high calcium scores and can thus be considered as

a demanding stress test for a coronary segmentation algorithm. In contrast, our larger internal database is more representative of the average results one could expect with modern 64-slices cardiac CTA.

The Rotterdam evaluation also allows the direct comparison with other existing methods. In Table 7.3, we report average results from other interactive methods. Our algorithm brings a noticeable robustness improvement (97.6% against 91.9% OV) over the method from Metz et al. (2008a) which implements a classical, two-point spatial minimal path technique based on image intensity and a Hessian-based vesselness feature. The method of (Szymczak, 2008) also relies on minimal paths. Instead of optimizing paths on a regular grid however, Szymczak (2008) rely on a so-called *core* graph, constructed from a pre-selection of spatial locations based on local intensity and topological criteria. To date, the only publicly ranked, interactive method to outperform our approach in terms of robustness is Friman et al. (2008b), an adaptation of Friman et al. (2008a) supplemented with minimal paths, which obtained $\texttt{OV} = 98.5\%$. It is worth noticing that this method required more interaction than ours, with the use of intermediate seed points. In contrast, we strictly limited ourselves to the provided start and end points.

7.7.2.2 Fully Automatic Workflow

We similarly evaluated our fully automatic approach. Results of the internal evaluation are given in Table 7.4. Understandably, automation degrades the results. We obtain an average overlap measure of OV=85%. One can notice a relatively high percentage of false positives (16.5%), explained by two factors. First, as we mentioned it in the qualitative analysis, the automatic algorithm was often able to extract actual branches or faint distal ends that were neglected from the manual delineation. Second, leakage remains an issue. In order to obtain truly reliable statistics, one should carefully complete the ground truth database to include all branches, even the smallest, faintest ones, up to their very distal end. This task would be particularly tedious and not necessarily of prime clinical relevance. More importantly, one should note the overlap result obtained for the main vessels (diameter greater than 1.5mm). The OT percentage of 92.9% is satisfyingly high, meaning that the algorithm is able, in a vast majority of cases, to extract all the principal coronary segments. Finally, accuracy results are on par with the measures obtained for the interactive workflow.

Table 7.5 summarizes the results for the Rotterdam evaluation. We emphasize that the Rotterdam framework focuses only on a subset (4 branches) of the coronary network, whereas our internal validation took into account both trees in their entirety. This explains, in particular, that the overlap results are slightly better for the Rotterdam evaluation. The overlap result OV of 91.3% is of the order the intra-observer variability (score of 50.9). Among publicly ranked fully automatic methods (see Table 7.6), the only algorithm to outperform our approach in terms of overlap is (Bauer and Bischof, 2008a), with an OV percentage of 92.7%. The method of (Bauer and Bischof, 2008a) employs a global approach, based on detection and graph-based reconnection of vessel segments, whereas our approach relies on a more local tracking scheme. Other automatic tracking schemes showing good performance are (Zambal et al., 2008) and (Tek et al., 2008), which both obtained average overlap measures of 84.7%. The approach of (Tek et al., 2008), based on (Gulsun and Tek, 2008b) which we previously mentioned, is more particularly related to our work. Besides our Bayesian formulation, the main differences are the incorporation of the radius dimension in the minimal path optimization, along with the use of the radius

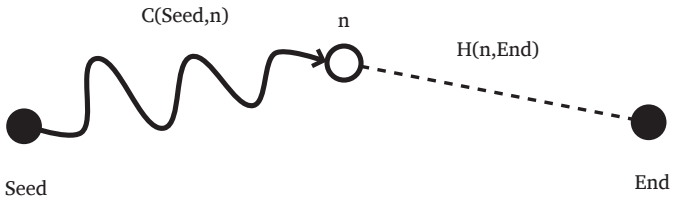


Figure 7.32: A^* optimization. Paths to a node n are ordered with respect to $C(Seed, n) + H(n, End)$, where $C(Seed, n)$ is the classical cumulative cost from the seed to n, and $H(n, End)$ is a heuristic estimation of the remaining cost from n to the end.

prior to guide the scale-space exploration.

7.8 Discussion and Perspectives

In this chapter, we have presented our first extraction strategy, based on discrete minimum path optimization. Our algorithm relies on the propagation of realizations of our geometric model as paths on a 4D space+scale graph. Among key design issues, we discussed the adequate discretization of the search space and the definition of the cumulative cost metric, which we derived from our seminal Bayesian model. Our cost formulation includes data likelihood and model-based prior terms which are learned automatically from our database. In this regard, we believe it improves on existing, more parameter-dependent techniques. Additionally, we introduced algorithmic refinements such as heuristic schemes limiting scale-space exploration. These schemes improve the computational and memory efficiency of our algorithm, making it usable on today's commercial workstations. Our qualitative and quantitative evaluation has demonstrated the robustness of our approach.

Among possible algorithmic extensions of the two-point workflow, the most immediate one is the implementation of a bidirectional propagation scheme, *i.e.*, propagating simultaneously from the seed *and* end point. Bidirectional propagation was shown to considerably reduce the extent of required exploration while yielding the exact same result as the one-seed propagation in the general minimal path problem (Deschamps and Cohen, 2001; Olabarriaga et al., 2003; Wink et al., 2000b) [4]. Propagating from the end point would correspond to a backward derivation of our Bayesian model. Thus, the radius transition prior $p(r_t|r_{t-1} = r^i_{t-1})$ would have to be reversed. The appropriate backward prior $p(r_{t-1}|r_t = r^i_t)$ should be extracted by re-integration from the learned joint density $\hat{f}(r_t, r_{t-1})$ (see Sec. 6.4.1 and Fig. 6.5).

Another well-spread refinement to minimal path techniques is the use of A^* heuristics to reduce the extent of graph exploration. The work of (Wink et al., 2000b) demonstrated their utility for vascular segmentation. As illustrated in Fig. 7.32, A^* algorithms rely on a heuristic estimation $H(n, End)$ of the remaining cost from a given node n to the End point. The idea is now to propagate hypotheses according to $C(Seed, n) + H(n, End)$. The propagation front is steered towards the end point by favoring the heuristically most promising hypotheses. To be *admissible*, *i.e.*, to still guarantee the optimality of the

4. In our case, as we dynamically limit the extent of scale-space exploration (by keeping only $H < |R|$ hypotheses per location), the exact result equivalence would not necessarily hold.

extracted path with respect to the cumulative metric C, the heuristic should always be an *under-estimation* of the actual remaining cost (Pearl, 1984). For instance, Wink et al. (2000b) use a simple admissible heuristic, defined as $H(c, End) = D(n, End) * U_{\min}$, with $D(n, End)$ an estimation of the distance to the end point (Manhattan distance for the L_1 norm) and $U_{\min}$ is the minimal update cost. It is straightforward to see that this heuristic is admissible, as no path could outperform it. It is worth noting that this heuristic would have no effect in our case, as we specifically designed our cost metric so that $U_{\min} = 0$ (see Sec. 7.2.3). We would always have $H(n, End) = 0, \forall n$. Designing tighter (more informative) yet still admissible heuristics might be delicate. Alternatively, one could look into inadmissible but reasonable heuristics, such as the active testing model from (Geman and Jedynak, 1996), reformulated as a A^* heuristic in (Yuille and Coughlan, 1997).

Fully automatic tree extraction could potentially be improved in several ways. First, one could aim at reducing false negatives (missed branches and continuations) by relaxing the stopping criterion and using techniques such as geodesic voting (Rouchdy and Cohen, 2008; Mille and Cohen, 2009) to extract true vascular segments *a posteriori* in a more robust manner. False positives (leakage) could be reduced thanks to pruning techniques (see (Gulsun and Tek, 2008a) for instance). We believe that post-processing through a more global approach, such as the graph-based reconnection/pruning scheme of (Bauer and Bischof, 2008a) would ease the disambiguation of situations such as leakage into the vein network.

The robustness and computational efficiency of our approach open promising perspectives for the improvement of clinical workflows. In particular, we believe that the two-point, interactive version of our algorithm can serve as a versatile basic tool to ease and speed up semi-automatic segmentation routines.

Chapter 8

Second Extraction Strategy: Tracking with Particle Filtering

In this chapter, we present our second extraction strategy based on particle filtering. It aims at the recursive estimation of the posterior *distribution* of our Bayesian model (Chapter 6) through the evolution of random samples, the so-called particles.

We first motivate the use of particle filters for vascular segmentation and present the core principles and tools of sequential Monte-Carlo methods. We then discuss our particular implementation, which includes a number of refinements, such as a dynamically adaptive propagation scheme, coupled with mean-shift clustering for branching detection and result extraction.

Compared to the minimal path method of Chapter 7, this extraction scheme does not enjoy the same workflow versatility. Particle filtering is only suited for direct, one-seed tracking of the coronary tree. On the other hand, it does not depend on the discretization of the search space and does not rely on heuristic selection schemes. Our evaluation demonstrates the practical potential of this approach in terms of robustness, accuracy and computational efficiency.

8.1 Problem Statement and Prior Art

In Chapter 6, we developed our Bayesian model and discussed the definition of its different components. Our problem of interest is the recursive estimation of the posterior distribution of our model, $p(x_{0:t}|z_{1:t})$. We recall the classical Bayesian recursion (Eq. 6.5):

$$p(x_{0:t}|z_{1:t}) = \frac{p(x_t|x_{t-1})p(z_t|x_t)}{p(z_t|z_{1:t-1})} p(x_{0:t-1}|z_{1:t-1}) = \frac{p(x_t|x_{t-1})p(z_t|x_t)}{\int p(z_t|x_t)p(x_t|z_{1:t-1})dx_t} p(x_{0:t-1}|z_{1:t-1}) \tag{8.1}$$

where $x_t, t \in [\![0, L]\!]$, correspond to the *states* of our model, and $z_t, t \in [\![0, L]\!]$, are sets of data *observations* available at each step t. The assumptions made in this derivation are that the states x_t follow a first order Markovian process and that observations z_t are conditionally independent given the model state x_t.

Other classical formulas in recursive Bayesian *filtering*[1] include the two-step recursion of the so-called marginal *filtering* distribution $p(x_t|z_{1:t})$. The first step is referred to as

1. Throughout this chapter, the term *filtering* is employed in its Bayesian sense, *i.e.*, the estimation of the posterior distribution $p(x_{0:t}|z_{1:t})$ given observations $z_{1:t}$ up to time t. *Smoothed* estimates, in contrast, are more intricate as they take *future* observations into account.

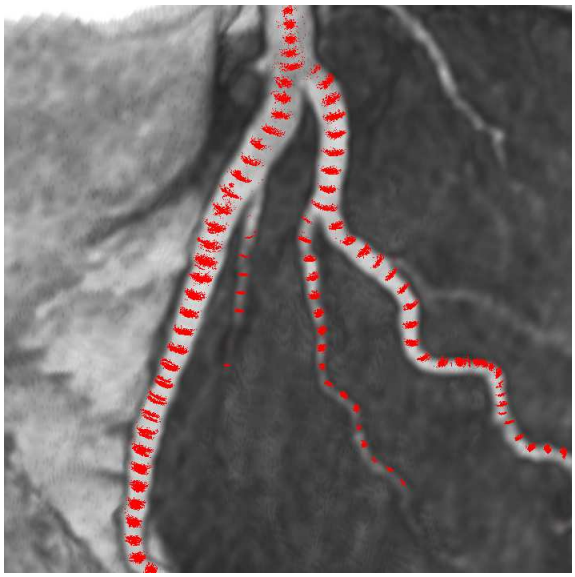


Figure 8.1: Particle filters for CTA coronary tracking. Generations of particles (red point clouds) are evolved iteratively to keep track of $p(x_{0:t}|z_{1:t})$, the posterior *distribution* of the vessel model $x_{0:t}$ given data observations.

the *prediction*:

$$p(x_t|z_{1:t-1}) = \int p(x_t|x_{t-1})p(x_{t-1}|z_{1:t-1})dx_t \tag{8.2}$$

This equation governs the transition of the marginal $p(x_{t-1}|z_{1:t-1})$ from step $t-1$ to step t, through the transition prior $p(x_t|x_{t-1})$. The second stage is the *update*, taking into account the new observations to obtain the filtering distribution at time t:

$$p(x_t|z_{1:t}) = \frac{p(z_t|x_t)p(x_t|z_{1:t-1})}{p(z_t|z_{t-1})} = \frac{p(z_t|x_t)p(x_t|z_{1:t-1})}{\int p(z_t|x_t)p(x_t|z_{1:t-1})dx_t} \tag{8.3}$$

As thoroughly discussed in (Doucet et al., 2000), Equations 8.1, 8.2 and 8.3 are deceptively simple. In general, the integrals involved cannot be derived analytically. Notable exceptions include Kalman filtering for linear Gaussian processes (Kalman, 1960) and grid-based methods (Arulampalam et al., 2002) for discrete state spaces. As was illustrated in Chapter 6, our process uses a non-linear observation model and non-Gaussian dynamics. In fact, as the different distributions composing our model are learned in a non-parametric manner, one cannot make any assumption about the overall process. Additionally, our 4D search space is too large for an adequate discrete optimization by grid-based methods to be computationally practical. Importantly, the extraction method presented in Chapter 7 indeed uses a discretization of the search space, but it relies on the empirical reformulation of the problem as the extraction of minimal paths and thus does not cope with the general problem of estimating the posterior distribution of our model. Its computational efficiency comes from its iterative, causal optimization scheme, which implicitly uses a selection rule (direct comparisons based on an empirical cumulative cost metric) to prune most of the hypotheses.

In order to deal with non-linear, non-Gaussian continuous state-space processes, it is then necessary to employ approximate techniques. Candidates include notably approximate grid-based techniques and extended Kalman filters (Arulampalam et al., 2002).

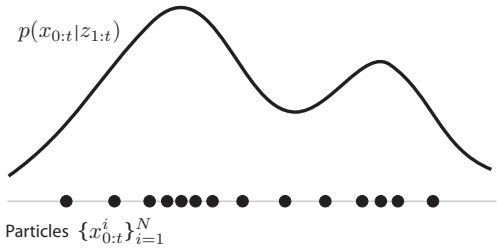


Figure 8.2: Monte-Carlo approximation of $p(x_{0:t|z_{1:t}})$ by a discrete set of samples $\{x^i_{0:t}\}^N_{i=1}$ identically distributed from $p(x_{0:t|z_{1:t}})$. The empirical approximation is given by: $\hat{p}(x_{0:t|z_{1:t}}) = \frac{1}{N}\sum_{i=1}^{N}\delta(x_{0:t} - x^i_{0:t})$.

Approximate grid-based techniques still require the discretization of the search space and potentially very costly resolution schemes. Extended Kalman filters rely on the local linearization of the process. One of their main drawback is still to estimate the posterior as a Gaussian and are thus unable to handle multi-modal posteriors. In our applicative context, this means that such filters cannot cope with bifurcations. An alternative, which we adopt in this work, is the use of particle filtering techniques.

Particle filtering techniques, also known as Sequential Monte-Carlo methods, rely on the stochastic evolution of a population of samples (the so-called particles). Their application to posterior filtering problems can be seen as a multi-hypothesis, stochastic tracking scheme, although we emphasize that their overall scope is much larger. Their application to CTA coronary tracking is illustrated in Fig. 8.1.

As was previously mentioned in Sec. 3.3.5, the use of particle filters for 3D vascular segmentation was first proposed in (Florin et al., 2005, 2006). This seminal work inspired several recent developments (Schaap et al., 2007a,b; Lesage et al., 2008; Allen et al., 2008). These techniques differ in their applicative scopes, their transition and observation models, their sampling schemes and proposed refinements (new sampling scheme in (Lesage et al., 2008), Markov Chain Monte-Carlo rejuvenation in (Allen et al., 2008)). These different works have demonstrated the attractiveness of this framework for vascular segmentation, in particular in terms of robustness to local perturbations (pathologies, image artifacts). Most deterministic tracking schemes (Kalman-based, for instance) only propagate limited information (moments of the posterior distribution) under very restrictive assumptions (linearity of state-space variations, in particular) which do not hold in practice. In contrast, particle filters approximate the entire posterior distribution through populations of samples, without any assumption of linearity or Gaussianity. They offer a generic, highly flexible framework. In this chapter, we show how they can be employed to develop our full Bayesian model and discuss the refinements we propose in our implementation.

8.2 Basics of Sequential Monte-Carlo Methods

8.2.1 Monte-Carlo Estimation

In the general, non-linear, non-Gaussian scenario, the main issue with the recursion equations of the previous section is the intractability of the potentially high-dimensional integrals involved. Monte-Carlo integration methods rely on empirical sampling to resolve

such issues (Doucet et al., 2001).

Assuming the possibility of simulating N random samples $\{x^i_{0:t}\}_{i=1}^N$ identically distributed from $p(x_{0:t}|z_{1:t})$ (*perfect Monte-Carlo sampling*), we obtain an empirical estimate $\hat{p}(x_{0:t}|z_{1:t})$ of $p(x_{0:t}|z_{1:t})$ as:

$$\hat{p}(x_{0:t}|z_{1:t}) = \frac{1}{N}\sum_{i=1}^{N}\delta(x_{0:t} - x^i_{0:t})$$

where $\delta(.)$ is the delta-Dirac measure. In other words, the distribution is approximated by a discrete sum of point-wise estimates at sampling locations (see Fig. 8.2). These samples $x^i_{0:t}$ are also referred to as the *particles*.

Additionally, one can easily estimate features of $p(x_{0:t}|z_{1:t})$ such as expectations of the form:

$$I(f_t) = \mathbb{E}_{p(x_{0:t}|z_{1:t})}[f_t(x_{0:t})] \triangleq \int f_t(x_{0:t})p(x_{0:t}|y_{1:t})dx_{0:t}$$

with f_t a function integrable with respect to $p(x_{0:t}|z_{1:t})$. For instance, for $f_t(x_{0:t}) = x_{0:t}$, $I(f_t)$ corresponds to the conditional mean of the process. A Monte-Carlo estimate $\hat{I}(f_t)$ of $I(f_t)$ is simply given by:

$$\hat{I}(f_t) = \frac{1}{N}\sum_{i=1}^{N} f_t(x^i_{0:t})$$

Such estimators are unbiased and converge almost surely with a rate of $O(1/N)$ as $N \to +\infty$ (Doucet et al., 2001). One fundamental observation is that the convergence rate is independent of the dimension of the state space to be estimated, making Monte-Carlo methods particularly attractive for complex estimation problems.

Perfect Monte-Carlo principles rely on the assumption that one can sample from $p(x_{0:t}|z_{1:t})$. Unfortunately, this is generally not possible, so alternative schemes must be devised. Markov Chain Monte-Carlo (MCMC) techniques offer a way to generate samples from such complex distributions (Gilks and Spiegelhalter, 1996), but the iterative procedures they employ make them impractical for recursive estimation. An alternative is the use of *importance sampling*.

8.2.2 Importance Sampling

Importance sampling relies on the use of a so-called *importance sampling distribution* $q(x_{0:t}|z_{1:t})$, also known as *proposal* distribution, to circumvent the issue of not being able to sample from the posterior distribution $p(x_{0:t}|z_{1:t})$ (Geweke, 1989).

In theory, very few restrictions are imposed on $q(x_{0:t}|z_{1:t})$. It may or may not depend on the observations $z_{1:t}$, as will be discussed in a subsequent section. Its support should, however, include that of $p(x_{0:t}|z_{1:t})$, and one should be able to easily sample from it.

The core idea of importance sampling is to obtain samples from $q(x_{0:t}|z_{1:t})$ and *correct* them with respect to $p(x_{0:t}|z_{1:t})$ (see Fig. 8.3). This correction is carried out through the introduction of *importance weighting*:

$$w(x_{0:t}) \propto \frac{p(x_{0:t}|z_{1:t})}{q(x_{0:t}|z_{1:t})} \tag{8.4}$$

Such a weighting function needs only to be determined up to a proportionality constant. From N particles $\{x^i_{0:t}\}_{i=1}^N$ identically distributed from $q(x_{0:t}|z_{1:t})$, a valid Monte-Carlo

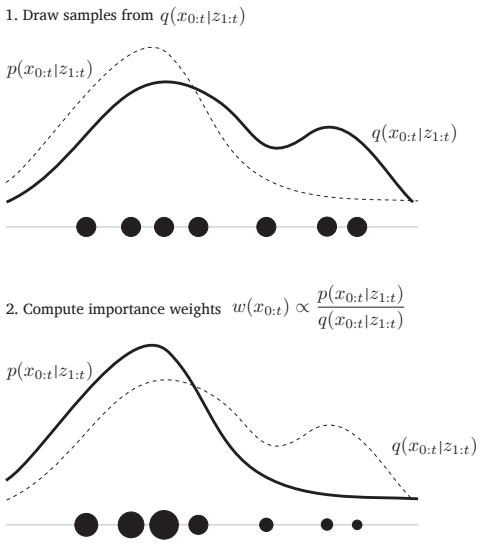


Figure 8.3: Importance Sampling. Samples from an *importance* distribution $q(x_{0:t}|z_{1:t})$ are weighted with respect to $p(x_{0:t}|z_{1:t})$ to obtain an approximation $\hat{p}(x_{0:t}|z_{1:t}) = \sum_{i=1}^{N} \tilde{w}_t^i \delta(x_{0:t} - x_{0:t}^i)$.

approximation of the posterior $p(x_{0:t}|z_{1:t})$ is now given by:

$$\hat{p}(x_{0:t}|z_{1:t}) = \sum_{i=1}^{N} \tilde{w}_t^i \delta(x_{0:t} - x_{0:t}^i) \tag{8.5}$$

where $\{\tilde{w}_t^i\}$ are the normalized importance weights associated to the particles:

$$\tilde{w}_t^i = \frac{w(x_{0:t}^i)}{\sum_{j=1}^{N} w(x_{0:t}^j)} \tag{8.6}$$

Similarly to perfect Monte-Carlo techniques, one can obtain estimates of expectations $I(f_t)$ as:

$$\hat{I}(f_t) = \sum_{i=1}^{N} f_t(x_{0:t}^i) \tilde{w}_t^i$$

Such estimators are biased for finite numbers of samples N but consistence and asymptotic convergence (in $O(1/N)$ for the bias and variance) are ensured by the satisfaction of a central limit theorem (Geweke, 1989).

8.2.3 Sequential Importance Sampling

Importance sampling is a general, powerful Monte-Carlo technique to circumvent the issue of estimating intractable integrals with distributions that cannot be directly sampled from. Its seminal formulation is however unsuited for recursive problems, as the weighting function (Eq. 8.4) involves states and data observations from the entire sequence. *Sequential importance sampling* consists in using importance sampling in a sequential, recursive

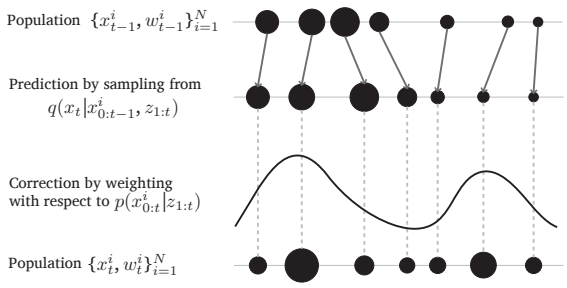


Figure 8.4: Sequential Importance Sampling (SIS). The new population $\{x^i_t, w^i_t\}^N_{i=1}$ is obtained by evolving $\{x^i_t, w^i_t\}^N_{i=1}$ by sampling from the importance distribution $q(x_t|x^i_{0:t-1}, z_{1:t})$ and correcting the previous weights $\{w^i_{t-1}\}$ (Eq. 8.8).

fashion, without modifying past samples. The aim is to keep the complexity of the sequential estimation process from increasing with the temporal length of the sequence. The key to achieve that goal is to consider an importance distribution that factorizes as follows:

$$q(x_{0:t}|z_{1:t}) = q(x_t|x_{0:t-1}, z_{1:t})q(x_{0:t-1}|z_{1:t-1}) \tag{8.7}$$

In other words, the importance distribution is now expressed recursively, with a prediction component $q(x_t|x_{0:t-1}, z_{1:t})$. For the sake of conciseness, we will, in the remaining of the document, refer to $q(x_t|x_{0:t-1}, z_{1:t})$ as the importance distribution.

This allows us to evaluate importance weights recursively (Doucet et al., 2001):

$$\begin{aligned} w^i_t &\propto \frac{p(x^i_{0:t}|z_{1:t})}{q(x^i_{0:t}|z_{1:t})} \\ &\propto \frac{p(z_t|x^i_t)p(x^i_t|x^i_{t-1})p(x^i_{0:t-1}|z_{1:t-1})}{q(x^i_t|x^i_{0:t-1}, z_{1:t})q(x^i_{0:t-1}|z_{1:t-1})} \\ &\propto w^i_{t-1}\frac{p(z_t|x^i_t)p(x^i_t|x^i_{t-1})}{q(x^i_t|x^i_{0:t-1}, z_{1:t})} \end{aligned} \tag{8.8}$$

One can note that it is only required to be able to evaluate the marginal likelihood $p(z_t|x^i_t)$ and transition prior $p(x^i_t|x^i_{t-1})$ in a point-wise manner, at sampled locations.

We can now describe the first, simplest particle filtering algorithm, known as Sequential Importance Sampling (SIS) filter. Its principle is illustrated in Fig. 8.4, with corresponding pseudo-code given in Fig. 8.5. At each iteration t, the previous population of weighted samples $\{x^i_{t-1}, \tilde{w}^i_{t-1}\}^N_{i=1})$ is evolved by sampling new states from the importance distribution (line 3) and updating associated weights following Eq. 8.8 (line 4). One can note that there is no need for normalized weights in this procedure. The SIS algorithm is particularly simple, but unfortunately generally inefficient in practice. As discussed in the next section, naive sequential importance sampling suffers from severe *degeneracy* issues as the length of the sequence increases.

8.2.4 Key Issues and Existing Refinements

Degeneracy issues The key issue with the SIS approach is that the variance of the importance weights can be shown to increase exponentially with the length of the sequence

SIS($\{x^i_{t-1}, w^i_{t-1}\}_{i=1}^N$)

1 **for** $i \leftarrow 1$ **to** N

2 **do**

▷ Prediction (sampling from the importance distribution)

3 $x^i_t \sim q(x_t|x^i_{0:t-1}, z_{1:t})$

▷ Update (computation of the weights from Eq. 8.8)

4 $w^i_t \leftarrow w^i_{t-1} \frac{p(z_t|x^i_t)p(x^i_t|x^i_{t-1})}{q(x^i_t|x^i_{0:t-1}, z_{1:t})}$

Figure 8.5: Sequential Importance Sampling (SIS) particle filter. Obtention of the weighted particle population $\{x^i_t, w^i_t\}_{i=1}^N$ from $\{x^i_{t-1}, w^i_{t-1}\}_{i=1}^N$ (single iteration).

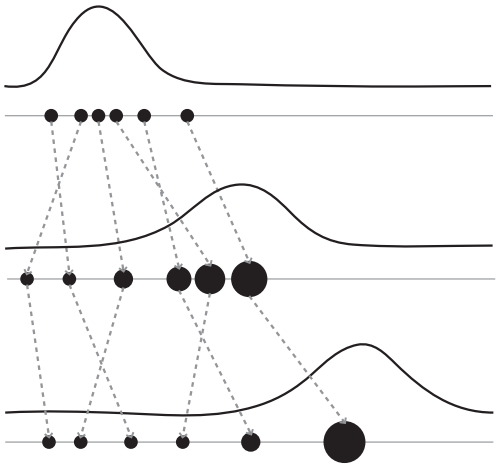


Figure 8.6: Degeneracy of the SIS filter. Using simple SIS filtering, the variance of importance weights can only increase over time. After a few iterations, only a small fraction of the particles have significant weights.

(Gilks and Spiegelhalter, 1996). In practice, this means that the population of samples quickly *degenerates* so that very few particles have significant weights (see Fig. 8.6). In other words, after a few iterations, the allocation of the particles becomes particularly inefficient, with a vast majority of samples lying in areas of very low probability mass. For tracking applications, this results in the target being quickly lost.

Degeneracy can be measured by the so-called effective sample size (ESS) (Kong et al., 1994; Liu and Chen, 1998). The ESS evaluates the quality of the approximation of the filtered distribution $p(x_t|z_{1:t})$ by the N particles sampled from the importance distribution $q(x_t|x^i_{0:t-1}, z_{1:t})$:

$$ESS(N) = \frac{N}{1 + d(p, q)}$$

with $d(q, p)$ the χ^2 distance between $p(x_t|z_{1:t})$ and $q(x_t|x_{0:t-1}, z_{1:t})$. This distance corresponds to the variance of the "true" weights $\mathring{w}^i(x_t) = p(x^i_t|z_{1:t})/q(x^i_t|x^i_{0:t-1}, z_{1:t})$:

$$d(p, q) = \text{var}\{\mathring{w}^i(x_t)\}$$

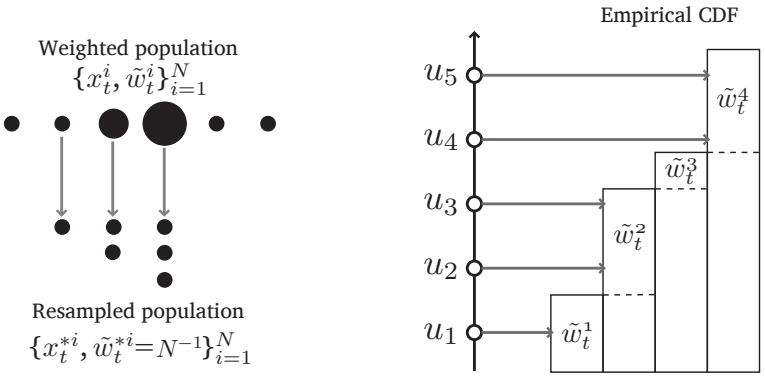


Figure 8.7: Resampling. Left: illustration of the general principle. Particles $\{x_t^{*i}, \tilde{w}_t^{*i}\}_{i=1}^N$ are selected from $\{x_t^i, \tilde{w}_t^i\}_{i=1}^N$ proportionally to the weights $\{\tilde{w}_t^i\}$. Some particles may be duplicated, others discarded. Note that the resampled weights $\{\tilde{w}_t^{*i}\}$ are all equal to N^{-1}. Right: systematic resampling (see text and Fig. 8.8). Particles x_t^2 and x_t^4 are selected twice, x_t^3 is discarded.

In practice, this variance cannot be evaluated, but can be estimated empirically thanks to the coefficient of variation of the weights w_t^i. We obtain:

$$\hat{ESS}(N) = \frac{1}{\sum_{i=1}^N (\tilde{w}_t^i)^2} \tag{8.9}$$

with $\{\tilde{w}_t^i\}$ the classical weights (Eq. 8.8) after normalization. Intuitively, the ESS gives the quality equivalence between the samples from the importance distribution and the true filtered distribution. The N samples from the importance distribution are, in terms of quality of the estimation, equivalent to $ESS(N)$ samples from the target distribution. First, one can note that one always has $ESS(N) < N$. The farther the importance distribution is from the true posterior, the faster the variance of the weights will increase and the faster $ESS(N)$ will drop, indicating severe degeneracy.

Less popular alternatives to the ESS include entropy and closely related Kullback-Leibler criteria (Fox, 2003; Soto, 2005; Lanz, 2007). They follow similar principles, based on measures of the disagreement between $p(x_t|z_{1:t})$ and $q(x_t^i|x_{0:t-1}^i, z_{1:t})$ to valuate degeneracy.

Several strategies can be devised and combined to cope with degeneracy. A classical approach is the use of *resampling* schemes. A more intricate one is the careful design of good importance distributions.

Resampling Resampling (partially) copes with degeneracy by reallocating the particles more efficiently. The general principle of resampling schemes is to obtain samples approximately distributed from the posterior distribution by (re-)sampling its weighted, importance sampling representation. To do so, samples x_t^i are selected with probability $\hat{w}_t^i$, that is, their normalized weight (see Fig. 8.7 left). This way, particles with low weights are likely to be discarded, and those with high weights to be duplicated.

Different resampling schemes have been proposed in the literature, among which multinomial resampling (Gordon et al., 1993), residual and stratified resampling (Higuchi, 1997;

RESAMPLE($\{x_t^i, \tilde{w}_t^i\}_{i=1}^N$)

▷ Build the empirical CDF from the normalized weights

1 $CDF[1] \leftarrow 0$
2 **for** $i \leftarrow 2$ **to** N
3 **do**
4 $CDF[i] \leftarrow CDF[i-1] + \tilde{w}_t^i$

▷ Draw the starting point

5 $u_1 \sim \mathbb{U}[0, N^{-1}]$

▷ Iterate on the CDF (index j)

6 $j \leftarrow 1$
7 **for** $i \leftarrow 1$ **to** N
8 **do**
9 $u_i \leftarrow u_1 + N^{-1}(i-1)$
10 **while** $u_i > CDF[j]$
11 **do**
12 $j \leftarrow j + 1$

▷ Assign

13 $x_t^{*i} \leftarrow x_t^j$
14 $\tilde{w}_t^{*i} \leftarrow N^{-1}$

Figure 8.8: Systematic Resampling. Resampling of a population $\{x_t^{*i}, \tilde{w}_t^{*i}\}_{i=1}^N$ from $\{x_t^i, \tilde{w}_t^i\}_{i=1}^N$.

Liu and Chen, 1998; Carpenter et al., 1999) and systematic resampling (Kitagawa, 1996; Carpenter et al., 1999). The reader may refer to (Arulampalam et al., 2002; Douc et al., 2005) for general overviews of these different techniques. Systematic resampling is certainly the most popular approach due to its simplicity and computational efficiency (in $O(N)$, with N the number of resampled particles). Moreover, it generally empirically outperforms the other methods (Arulampalam et al., 2002; Douc et al., 2005). The principle of systematic resampling is illustrated in Fig. 8.7, right. Corresponding pseudo-code is given in Fig. 8.8. The procedure starts by building the empirical CDF from the normalized importance weights. Then, a single random draw $u_1 \sim \mathbb{U}[0, N^{-1}]$ is performed and serves as the starting point (line 5). The value u_i corresponds to the CDF value for which the i^{th} resampled particle will be selected. Each u_i is evenly spaced as $u_i - u_{i-1} = N^{-1}$ (line 9). One should note that after resampling, weights $\{w_t^{*i}\}$ are uniform (line 14). The resampled particles $\{x^{*i}\}$ are indeed approximately identically distributed with respect to the empirical posterior. Finally, we emphasize that for the sake of clarity, we only considered the case of a fixed number N of samples. The procedure can be straightforwardly adapted to a variable number $N_t^* \neq N_t$ of resampled particles.

Coupling the SIS algorithm (Fig. 8.5) with a resampling step, one obtains the generic Sampling-Importance-Resampling (SIR) algorithm, also known as Bootstrap (Gordon et al., 1993) or Condensation filter (Isard and Blake, 1998b) [2]. Resampling is a key ingre-

2. As discussed thereafter, authors in (Gordon et al., 1993; Isard and Blake, 1998b) only considered the case of the transition prior $p(x_t|x_{t-1})$ as an importance distribution, although resampling principles apply naturally to any importance distribution.

dient to the practical usability of relatively simple, generic particle filtering techniques on a wide array of applications.

It can be shown that resampling injects additional Monte-Carlo variation, which may deteriorate the performance of the estimator (Chopin, 2004). If some well-spread techniques such as the Sampling-Importance-Resampling (SIR) / Bootstrap / Condensation algorithms (Gordon et al., 1993; Isard and Blake, 1998b) implement resampling at each step, it might be advisable to resample only when some degeneracy measure (ESS, for instance) drops below a certain threshold (Liu and Chen, 1998).

Sample impoverishment Although resampling schemes effectively reallocate samples, they can also lead to a loss of diversity among resampled particles, with a few dominant particles selected many times. This phenomenon is particularly acute for processes with low observation noise. A first partial remedy is to limit resampling steps, as stated previously. A second proposal is the use of regularization (kernel density estimation) to perturbate the samples, as done in the so-called regularized particle filter (Musso et al., 2001). Another, more formal approach is the implementation of Markov Chain Monte-Carlo (MCMC) moves (Gilks and Spiegelhalter, 1996; Gilks and Berzuini, 2001) to "rejuvenate" the population. This approach was recently applied to vascular segmentation in (Allen et al., 2008), with the aim of increasing the robustness of bifurcation detection.

Choice of importance distribution A key design issue for particle filters is the choice of an adequate importance distribution. This distribution, which serves a sampling basis for the prediction step of the filter (see the SIS procedure, Fig. 8.5), conditions critically the behavior of the filter. A good importance distribution, *i.e.*, close to the target filtering distribution, considerably alleviates most practical issues, including degeneracy and sample impoverishment.

In general, one will employ a (prediction) importance distribution which depends only on x_{t-1} and possibly on z_t, so that $q(x_t|x^i_{0:t-1}, z_{1:t}) = q(x_t|x^i_{t-1}, z_t)$. The optimal importance distribution $q_{opt}(x_t|x^i_{t-1}, z_t)$, minimizing the variance of the true weights $\mathring{w}^i_t$, is $q_{opt}(x_t|x^i_{t-1}, z_t) = p(x_t|x^i_{t-1}, z_t)$ (Doucet et al., 2000). Unfortunately, it is generally impossible to sample directly from this distribution (Arulampalam et al., 2002).

In practice, the design of a good importance distribution can be intricate. A general advice is that $q(x_t|x^i_{t-1}, z_t)$ should be slightly larger than the marginal $p(x_t|x^i_{t-1}, z_t)$ so as to ensure good tracking robustness. A too tight importance distribution may lead to a premature loss of the target. On the other hand, choosing a too large importance distribution is detrimental to the computational efficiency. If the prediction is constrained only loosely, a high number of particles will be needed to obtain good estimates.

A particularly popular choice of importance distribution is the use of the transition prior (Gordon et al., 1993; Isard and Blake, 1998b). Setting $q(x_t|x^i_{t-1}, z_t) = p(x_t|x^i_{t-1})$, the weight update (Eq. 8.8) simplifies as:

$$\begin{aligned} w^i_t &\propto w^i_{t-1} \frac{p(z_t|x^i_t)p(x^i_t|x^i_{t-1})}{q(x^i_t|x^i_{t-1}, z_t)} \\ &\propto w^i_{t-1} p(z_t|x^i_t) \end{aligned} \tag{8.10}$$

Weight update then consists solely of a correction by the likelihood $p(z_t|x^i_t)$. An important remark is that this importance distribution indeed does not take into account the most recent observations z_t. This may be particularly problematic in case of strong disagreement

between the prior and the likelihood (situations difficult to predict based solely on the transition prior).

Among refinements aiming at designing better importance distributions, one can cite Rao-Blackwellization (Casella and Robert, 1996), suitable when part of the state-space follows a linear model, and the use of the Unscented transform (Van der Merwe et al., 2001), constructing a Gaussian approximation of the optimal importance distribution $p(x_t|x_{t-1}^i, z_t)$. The use of bridging densities (Godsill and Clapp, 2001) or progressive correction schemes (Oudjane and Musso, 2000) can help coping with suboptimal proposals. Another popular approach forms the family of *auxiliary* particle filters (Pitt and Shephard, 1999), which we employ in this work and discuss in Sec. 8.3.

In a previous work (Lesage et al., 2008) (see Appendix E), we showed how one can build a discrete approximation of the posterior distribution, taking into account the latest available observations. We demonstrated that sampling from an appropriate importance distribution considerably reduces the number of samples needed to obtain a good estimation performance. The sampling scheme suffers from two main drawbacks. First, it implies a discretization of the search space, which may impact the accuracy of the estimations negatively. Second, the construction of the full posterior, even discretized, can be relatively costly. In (Lesage et al., 2008), we used a voxel-wise likelihood model relying notably on mathematical morphology. Because of the simplicity of the evaluations involved, the algorithm maintained a good level of computational efficiency. On the other hand, the use of a low-level likelihood model made the technique ill-adapted to the robust tracking of the smallest, low-contrast arteries.

Adaptation of the number of particles A last, noteworthy refinement to classical particle filtering is the adaptation of the number of samples. So far, we only considered the number of particles as a fixed parameter N of the filter. A general observation is that the higher this number, the better the estimation (convergence in $O(1/N)$). In practice, the number of particles required to obtain "good" estimates depends notably on the difficulty and dimensionality of the problem at hand. In a vast majority of cases, N is thus set empirically as a tradeoff between robustness, accuracy and computational cost. Considerations on Cramer-Rao bounds can help set N in a less heuristic manner (Simandl and Straka, 2002).

The number of particles can also be adapted *dynamically*, considering a time-dependent number N_t. The idea is to allocate N_t more efficiently. Less particles may suffice for easy iterations, while more intricate situations may require more samples. The goal is thus twofold: increase the quality of the estimates and potentially improve the computational efficiency of the filter. The reader may note that the idea of "intricate" versus "easy" situations closely relate to the aforementioned notion of degeneracy. The work of (Fox, 2003) proposes to adapt the number of particles as a function of the Kullback-Leibler (KL) distance between the marginal filtered posterior and the proposal. This work was extended in (Soto, 2005) with more refined estimation schemes of the KL distance. Lanz (2007) developed an entropy-based criterion. The work of (Koller and Fratkina, 1998) relies on a heuristic, point-wise quality criterion: keeping a fixed sum of particle likelihoods. The general idea is to augment the number N_t if current samples are of low quality (*i.e.*, with low weights). Straka and Simandl (2004) elaborate on this principle, and, in a subsequent work (Straka and Simandl, 2006), propose to allocate the number of particles as a function of a *fixed*, pre-defined effective sample size (ESS). We reuse that idea in our work, as discussed further in Sec. 8.3.

8.3 Adaptive Particle Filtering

Our implementation includes a couple refinements compared to textbook particle filters such as SIR and Bootstrap technique. First, our algorithm relies on an adaptive scheme, known as *auxiliary* sampling, to construct the proposal distribution. Second, it uses a time-varying number of samples, dynamically allocated with respect to the difficulty of the local estimation.

8.3.1 Adaptation of the Proposal: Auxiliary Particle Filters

The principle of *auxiliary* particle filtering (APF) (Pitt and Shephard, 1999) is to construct the importance distribution empirically, by simulation, and take into account the current observations z_t. A particle filter taking into account the new observations in the proposal is said to be *adapted*. Auxiliary sampling is a two-pass process, where the first *simulation* pass allows a one-step look-ahead to improve the effective allocation of the particles in the second pass.

Mathematically, APF is more easily described with the help of a discrete auxiliary variable, the index k, which encodes backward parental links between the different samples. APF relies on the construction of an importance distribution $q(x_t, k|z_{1:t})$ to sample couples $\{x_t^i, k^i\}_{i=1}^N$, with k^i the index of the particle at time $t-1$. The distribution $q(x_t, k|z_{1:t})$ is constructed to approach $p(x_t, k|z_{1:t})$. If one can obtain samples $\{x_t^i, k^i\}$ approximately from $p(x_t, k|z_{1:t})$, one can simply drop the indices k^i. The population $\{x_t^i\}$ would then correspond to valid samples from $p(x_t|z_{1:t})$, the marginalized *posterior*.

Bayes' rule gives us the development of $p(x_t, k|z_{1:t})$:

$$\begin{aligned} p(x_t, k|z_{1:t}) &\propto p(z_t|x_t)p(x_t, k|z_{1:t-1}) \\ &\propto p(z_t|x_t)p(x_t|k, z_{1:t-1})p(k|z_{1:t-1}) \\ &\propto p(z_t|x_t)p(x_t|x_{t-1}^k)\tilde{w}_{t-1}^k \end{aligned} \tag{8.11}$$

This proportionality equation involves the transition distribution $p(x_t|x_{t-1}^k)$ and the weight w_{t-1}^k of the k^{th} particle x_{t-1}^k, along with the likelihood distribution $p(z_t|x_t)$ over all states. The importance distribution $q(x_t, k|z_{1:t})$ is decomposed as:

$$q(x_t, k|z_{1:t}) = q(k|z_{1:t})q(x_t|k, z_{1:t}) \tag{8.12}$$

The two terms are set as:

$$q(k|z_{1:t}) \propto p(z_t|\mu_t^k)\tilde{w}_{t-1}^k \tag{8.13}$$

$$q(x_t|k, z_{1:t}) \triangleq p(x_t|x_{t-1}^k) \tag{8.14}$$

so that:

$$q(x_t, k|z_{1:t}) \propto p(z_t|\mu_t^k)p(x_t|x_{t-1}^k)\tilde{w}_{t-1}^k \tag{8.15}$$

where μ_t^k is some statistical characteristics of $p(x_t|x_{t-1}^k)$, for instance the mean state $\mathbb{E}[x_t|x_{t-1}^i]$. In our work, we consider simply a sample from the transition prior, with $\mu_t^k \sim p(x_t|x_{t-1}^i)$. This approach is similar to (Florin et al., 2006) and (Soto, 2005), who also implemented that particular case of APF and demonstrated its efficiency.

Comparing Eq. 8.11 and 8.15, on can note that the main difference is, in the importance distribution, the use of point-wise estimates μ_t^k. The samples $\{x_t^i, k^i\}$ obtained from $q(x_t, k|z_{1:t})$ are thus assigned importance weights as:

$$\begin{aligned} w_t^i &\propto w_{t-1}^{k^i} \frac{p(z_t|x_t^i)p(x_t^i|x_{t-1}^{k^i})}{q(x_t^i, k^i|z_{1:t})} \\ &= \frac{p(z_t|x_t^i)}{p(z_k|\mu_t^{k^i})} \end{aligned} \tag{8.16}$$

The algorithmic description of the APF is given in Fig. 8.9. It can be decomposed in three main steps:

1. construction of the empirical importance distribution $q(k|z_{1:t})$:
   - simulation: sampling of the population $\{\mu_t^i, w_t^i\}$ and computation of their associated weights w_t^i (line 3);
2. empirical sampling from $q(k|z_{1:t})$:
   - resampling from $\{\mu_t^i, w_t^i\}$ to obtain $\{k^i\}$ (line 8);
3. sampling of the new population $\{x_t^i, \tilde{w}_t^i\}$ from $q(x_t|k, z_{1:t})$:
   - prediction of new states from $\{x_{t-1}^{k^i}\}$ (line 11);
   - weights update following Eq. 8.16 (line 12).

The simulated population $\{\mu_t^i, w_t^i\}$ is used as an empirical representation of $q(k|z_{1:t})$. Classical resampling (Fig. 8.8) is employed to obtain the population of indices $\{k^i\}$. If μ_t^i is obtained as a sample from $p(x_t|x_{t-1}^i)$, one iteration of APF corresponds two consecutives SIS (Fig. 8.5) passes with intermediate resampling.

Another interpretation is that APF first performs a selection with replacement on the population at $t-1$, with respect to an empirical approximation of $p(x_{t-1}|z_{1:t})$ (one-step look-ahead). It then evolves the population classically through the transition prior $p(x_t|x_{t-1})$.

Providing that the point-wise estimates $\{\mu_t^i\}$ give a good characterization of $p(x_t|x_{t-1}^i)$ (low noise process), samples generated by the APF are likely to be well located, close to the true state x_t. Compared to classical SIS/SIR with the prior as importance distribution, the two-pass design of the APF generally reduces degeneracy issues, decreasing the variance of the weights by incorporating information on the most recent observations. A counterpart of APF is its higher computational complexity as it actually requires the evaluation of $2N$ particles per iteration.

8.3.2 Dynamic Adaptation of the Number of Particles

As was previously mentioned, a possible refinement of standard particle filtering is to vary *dynamically* the number of samples. We now consider a varying number N_t of particles to be determined at each iteration. Among the existing techniques cited in Sec. 8.2.4, we reuse the idea of (Straka and Simandl, 2006) for its simplicity and intuitiveness.

Straka and Simandl (2006) proposed to choose N_t so as to attain a *fixed* effective sample size (ESS). The ESS, discussed in Sec. 8.2.4, corresponds to a measure of the degree of degeneracy affecting the filter. In other words, targeting a fixed ESS N^* aims at maintaining a stable quality of the estimation. In order to attain N^*, a relatively low number of particles N_t will be required in situations where the importance distribution coincides closely with the filtered posterior (low variance of the weights). Conversely,

APF($\{x_{t-1}^i, \tilde{w}_{t-1}^i\}_{i=1}^N$)

▷ Simulation

for $i \leftarrow 1$ **to** N

do

obtain μ_t^i

$w_t^i \leftarrow p(z_t|\mu_t^i)\tilde{w}_{t-1}^i$

▷ Normalization

for $i \leftarrow 1$ **to** N

do

$w_t^i \leftarrow \frac{w_t^i}{\sum_{j=1}^N w_t^j}$

▷ Sampling of the indices by resampling (empirical sampling from $q(k|z_{1:t})$)

$\{k^i\}_{i=1}^N \leftarrow$ RESAMPLE($\{\mu_t^i, w_t^i\}_{i=1}^N$)

▷ Sampling of the new particles from $q(x_t|k, z_{1:t})$

for $i \leftarrow 1$ **to** N

do

$x_t^i \sim q(x_t|k^i, z_{1:t}) = p(x_t|x_{t-1}^{k^i})$

$w_t^i \leftarrow \frac{p(z_t|x_t^i)}{p(z_k|\mu_t^{k^i})}$ ▷ Eq. 8.16

▷ Normalization

for $i \leftarrow 1$ **to** N

do

$\tilde{w}_t^i \leftarrow \frac{w_t^i}{\sum_{j=1}^N w_t^j}$

Figure 8.9: Auxiliary particle filter (APF). Obtention of the weighted particle population $\{x_t^i, \tilde{w}_t^i\}_{i=1}^N$ from $\{x_{t-1}^i, \tilde{w}_{t-1}^i\}_{i=1}^N$ (single iteration).

a larger N_t will be needed in more difficult cases, such as noisy, low-contrast vessels, bifurcations and pathological cases.

We recall that the ESS of a population of N particles is defined as:

$$ESS(N) = \frac{N}{1 + d(p, q)} \tag{8.17}$$

Fixing a target ESS N^*, we have:

$$N_t = \lceil N^*(1 + d(p, q)) \rceil \tag{8.18}$$

The quantity $d(q, p)$, which is the χ^2 distance between the marginal posterior $p(x_t|z_{1:t})$ and the proposal $q(x_t|x_{t-1}, z_t)$, is not available analytically and has to be integrated numerically.

Assuming the availability of an *initial* population $\{x_t^i, \tilde{w}_t^i\}_{i=1}^{N_t^1}$ of N_t^1 samples from $q(x_t|x_{t-1}, z_t)$ with associated (normalized) importance weights, $d(q, p)$ can be estimated empirically by the coefficient of variation cv^2:

$$\hat{d}(p, q) = cv^2 = \frac{N_t^1}{\sum_{i=1}^{N_t^1} (\tilde{w}_t^i)^2} - 1 \tag{8.19}$$

By combining Eq. 8.19 and Eq. 8.18, we can derive the number of particles N_t as:

$$N_t = \lceil N^* \left(\frac{N_t^1}{\sum_{i=1}^{N_t^1} (\tilde{w}_t^i)^2} \right) \rceil \tag{8.20}$$

One always has $N_t \geq N^*$.

In (Straka and Simandl, 2006), a slightly more intricate estimation scheme is employed. We chose to rely on the estimation of $d(q,p)$ by cv^2 for the sake of simplicity and computational efficiency. The major issue with this scheme is the need for an initial population of weighted samples $\{x_t^i, \tilde{w}_t^i\}_{i=1}^{N_t^1}$. Existing sample size adaptation techniques (Koller and Fratkina, 1998; Fox, 2003; Soto, 2005; Straka and Simandl, 2006) advocate the implementation of a *burn-in* phase at each iteration. For instance, N_t can be estimated through the N^* first samples, leaving $N_t - N^*$ additional particles to be drawn. In our work, we propose to combine auxiliary particle filtering (APF) with an adaptive number of particles, using the simulation phase of the APF to estimate N_t.

8.3.3 Auxiliary Particle Filter with Adaptive Number of Particles (AAPF)

In Sec. 8.3.1, we described the principles of auxiliary particle filtering (APF), which relies on a simulation pass to construct an empirical proposal distribution taking into account most recent observations. In Sec. 8.3.2, we discussed how the number of allocated samples can be adjusted dynamically in function of the difficulty of the local estimation. We now show how both techniques can be combined in a natural fashion, yielding our adaptive auxiliary particle filtering algorithm (AAPF).

The estimation of the number of particles N_t requires an initial population of N_t^1 samples $\{x_t^i, \tilde{w}_t^i\}_{i=1}^{N_t^1}$. We propose to use the simulated population $\{\mu_t^i, w_t^i\}_{i=1}^{N_{t-1}}$ of the APF to evaluate N_t. The algorithmic description of our algorithm is given in Fig. 8.10. The main difference with classical APF (Fig. 8.9) is the estimation of N_t at line 8. We discuss the choice of target ESS N^* in Sec. 8.5 with respect to our application. This parameter intuitively balances overall estimation quality with computational cost.

From a theoretical point of view, it is important to note that the initial population $\{\mu_t^i, w_t^i\}_{i=1}^{N_{t-1}}$ on which N_t is estimated does not come from the APF proposal $q(x_t, k|z_{1:t})$, but from the transition prior used for the simulation (line 3). This means that the coefficient of variation of these initial weights is not an accurate estimation of $d(p,q)$, the χ^2 distance between the marginal posterior and the APF proposal. It is, however, a good measure of the complexity of the current iteration. If simulation by the transition prior leads to a large variance of these weights, chances are the APF will also have a relatively large weight variance. In practice, the APF generally reduces the weight variance by better allocating subsequent samples, so that this scheme tends to slightly overestimate N_t. In difficult situations, our technique reinforces the correction from APF by allocating more samples. As demonstrated in Sec. 8.5, this combination leads to a quantitative gain in robustness. In particular, it results in a more robust behavior at bifurcations, easing their detection (Sec. 8.4.3).

8.4 AAPF for CTA Coronary Segmentation

The AAPF algorithm described in the previous section is generic and could be applied to any Bayesian sequential estimation problem. We now discuss its particular use for our

AAPF($\{x_{t-1}^i, \tilde{w}_{t-1}^i\}_{i=1}^{N_{t-1}}$)

▷ Simulation

for $i \leftarrow 1$ **to** N_{t-1}

do

$\mu_t^k \sim p(x_t|x_{t-1}^i)$

$w_t^i \leftarrow p(z_t|\mu_t^i) w_{t-1}^i$

▷ Normalization

for $i \leftarrow 1$ **to** N_{t-1}

do

$w_t^i \leftarrow \frac{w_t^i}{\sum_{j=1}^{N_{t-1}} w_t^j}$

▷ Estimation of the number of particles N_t (Eq. 8.20)

$N_t \leftarrow \lceil N^*\left(\frac{N_{t-1}}{\left(\sum_{i=1}^{N_{t-1}} (w_t^i)^2\right)}\right) \rceil$

▷ Sampling of the indices by resampling (empirical sampling from $q(k|z_{1:t})$)

$\{k^i\}_{i=1}^{N_t} \leftarrow$ RESAMPLE($\{\mu_t^i, w_t^i\}_{i=1}^{N_{t-1}}$)

▷ Sampling of the new particles from $q(x_t|k, z_{1:t})$

for $i \leftarrow 1$ **to** N_t

do

$x_t^i \sim p(x_t|x_{t-1}^{k^i})$

$w_t^i \leftarrow \frac{p(z_t|x_t^i)}{p(z_k|\mu_t^{k^i})}$ ▷ Eq. 8.16

▷ Normalization

for $i \leftarrow 1$ **to** N_t

do

$\tilde{w}_t^i \leftarrow \frac{w_t^i}{\sum_{j=1}^{N_t} w_t^j}$

Figure 8.10: Auxiliary particle filter with adaptive number of particles (AAPF). Obtention of the weighted particle population $\{x_t^i, \tilde{w}_t^i\}_{i=1}^{N_t}$ from $\{x_{t-1}^i, \tilde{w}_{t-1}^i\}_{i=1}^{N_{t-1}}$ (single iteration).

application of choice, CTA coronary segmentation.

In our implementation, AAPF is used to estimate our *full* Bayesian model (Chapter 6). We recall that our state-space model is defined as $x_t = (p_t, d_t, r_t)$, with p_t the centerline position, d_t the tangential position and r_t the vessel radius (see Chapter 4). Particles $\{x_t^i\}$ evolved by the AAPF are thus state realizations of our model.

The workflow employed for our particle filtering approach is the same as for our full tree, automatic minimal path method (Sec. 7.6.2). From each of the detected ostia locations, we initialize our AAPF filter with a population of N_0 particles sampled uniformly among all possible directions and radiuses. We used $N_0 = 2000$ in our tests.

The AAPF algorithm involves the evaluation of likelihood terms (see Sec. 6.3) and sampling from transition priors (see Sec. 6.4). We recall that the different distributions are all learned in a non-parametric fashion from our ground-truth database (see Sec. 6.2). The learning process is dependent on the discretization resolution of the geometric model. We used a fixed step $s = 0.3mm$, of the order of the intra-slice resolution of our CTA data. Please note that this discretization step corresponds to the spacing of *control* points

Variable	discretization step	range	Number of values
Radius values: r_{t-1}, r_t	0.03mm	$[\![0.1, 3.97]\!]$	$N_r = 130$
`MFlux` responses: y_t	0.5H.U	$[\![-49.5, 150]\!]$	$N_y = 400$
Angular variation: α_t	$\frac{\pi}{198}$ rad	$[\![0, \frac{\pi}{2}]\!]$	$N_\alpha = 100$

Table 8.1: Parameters used for the discretization of the different distributions of our model.

$\|c_{t-1} - c_t\| = s, \forall t$ (see Chapter 4) and controls the spatial progression of the particle filter along the vessels.

In the following, we first discuss the evaluation of the likelihood terms (Sec. 8.4.1) and the sampling from the prior distributions (Sec. 8.4.2). As the AAPF propagates along the coronary tree, we use mean-shift clustering to explicitly detect bifurcations (Sec. 8.4.3). Each branch is evolved independently, until a stopping criterion is fulfilled (Sec. 8.4.4). A compact centerline+radius representation of the extracted tree is obtained thanks to the mean-shift modes (Sec. 8.4.5).

8.4.1 Evaluation of the Likelihood

We recall that the valuation of our likelihood model is decomposed as:

$$p(z_t|x_t = x_t^i) = \frac{p_v(y_t^i|r_t = r_t^i)}{p_{bg}(y_t^i)} \prod_{y_t^j \in z_t, \forall j} p_{bg}(y_t^j) \tag{8.21}$$

where observations at time t are $z_t = \{y_t^i\}_{i=1}^{N_t}$, with y_t^i an elementary observation obtained as the `MFlux` feature response for the state realization (particle) $x_t^i = (p_t^i, d_t^i, r_t^i)$. The distribution $p_v(y_t|r_t)$ values the *vessel* likelihood, *i.e.*, the probability that an observation y_t^i comes from an actual coronary in the image. Conversely, $p_{bg}(y_t^i)$ values the probability that the observation comes from the *background*.

Contrarily to the minimal path approach (Chapter 7), tracking with particle filters develops all considered realizations in parallel. The algorithm always deals with realizations of the same generation t, so that all the terms involved need only to be known up to a normalization constant. One can drop all the terms which are independent of the realization, such as the product of background likelihoods $\prod_{y_t^j \in z_t, \forall j} p_{bg}(y_t^j)$.

The distributions p_v and p_{bg} are extracted from non-parametric joint densities (see Sec. 6.2). We store them as histogram densities, discretized at relatively fine resolutions and directly integrated from the learned joint densities. The discretization parameters used in practice are given in Table 8.1. For instance, to encode the distributions $p_v(y_t|r_t = r_t^i)$ for the entire range of possible radiuses, we use $N_r = 130$ histograms of $N_y = 400$ discretized `MFlux` values. Our choices of discretization resolutions correspond to tradeoffs between memory requirements and the limitation of discretization effects. The reader may note for example that the radius discretization (resolution of 0.03mm) is still ten times finer than that used in our minimal path approach (Chapter 7). In practice, the histograms are used as simple lookup tables to value the likelihood terms. To avoid numerical instabilities in likelihood ratios $\frac{p_v(y_t^i|r_t=r_t^i)}{p_{bg}(y_t^i)}$, we impose a lower bound $p_{min} = 10^{-6}$ on $p_{bg}(y_t^i)$.

8.4.2 Sampling from the Prior Densities

Besides likelihood evaluation, AAPF requires the ability to sample from priors $p(x_t|x_{t-1} = x^i_{t-1})$. Our transition model (Sec. 6.4) is given by:

$$p(x_t|x^i_{t-1}) = p(r_t|r^i_{t-1})p(d_t|d^i_{t-1}, r^i_{t-1}) = p(r_t|r^i_{t-1})p(\alpha_t|r^i_{t-1}) \tag{8.22}$$

with $\alpha_t = \arccos d_t \cdot d_{t-1}$, the angular variation between d_{t-1} and d_t. The radius and angular transition priors are also learned non-parametrically. To be able to sample from these distributions, we first discretize them, similarly to the likelihood terms, and store them as histograms. Discretization parameters are given in Table 8.1. From these histogram representations, we obtain new radiuses and angular variations using multinomial sampling.

For a given particle $x^i_{t-1} = (p^i_{t-1}, d^i_{t-1}, r^i_{t-1})$, we obtain $x^i_t = (p^i_t, d^i_t, r^i_t)$ as follows:

1. new radius $r^i_t \sim \mathcal{H}(r_t|r^i_{t-1})$ (multinomial sampling from the radius prior histogram for r^i_{t-1});
2. new direction d^i_t:
   - obtain an orthonormal basis (d^i_{t-1}, v_1, v_2) (d^i_{t-1} as the x-axis by convention);
   - sample $\alpha^i_t \sim \mathcal{H}(\alpha_t|r^i_{t-1})$ (multinomial sampling from the angular prior histogram for r^i_{t-1});
   - sample $\beta^i_t \sim \mathbb{U}[0, 2\pi[$ (second rotation angle uniformly distributed);
   - obtain $d^i_t = cos(\alpha^i_t)d^i_{t-1} + sin(\alpha^i_t)(cos(\beta^i_t)v_1 + sin(\beta^i_t)v_2)$ (combination of the rotations);
3. new position $p^i_t = p^i_{t-1} + \frac{s(d^i_t + d^i_{t-1})}{2}$ (with $s = 0.3mm$ the spatial discretization step).

8.4.3 Detection of Bifurcations by Mean-Shift Clustering

With respect to the posterior distribution, bifurcations correspond to the apparition of multiple modes. Although particle filters are theoretically able to cope with multi-modal distributions, they raise practical issues in sequential estimation schemes. Because of the sample impoverishment induced by resampling procedures (see Sec. 8.2.4), secondary modes are quickly lost as all particles are re-allocated on the dominant mode. In our applicative context, this means that the filter will robustly track only one branch. As can be seen in Fig. 8.11 (left and middle), secondary branches are progressively depopulated and lost. These observations led previous works (Florin et al., 2005, 2006; Allen et al., 2008) to perform an explicit detection of the bifurcations by clustering techniques. As illustrated in Fig. 8.11 (right), the detection of the bifurcations allows for their independent propagation and re-population, so that secondary branches are preserved.

Compared to classical SIR implementations, our AAPF algorithm improves the behavior at bifurcations. As illustrated in Fig. 8.11, AAPF generally tracks secondary branches on a longer portion than SIR. More importantly, small, secondary branches are frequently missed by the SIR implementation. AAPF, on the other hand, consistently achieves a better sampling of asymmetric bifurcations. We explain this observation by a combination of auxiliary particle filtering and the dynamic adaptation of the number of samples. The number of particles selected by our scheme tends to be larger in bifurcation areas, as the apparition of multiple modes tends to increase the variance of the weights (see Fig. 8.12).

Although AAPF sensibly improves the algorithm's robustness at bifurcations, it still cannot track several branches in parallel until their distal ends. Secondary branches are

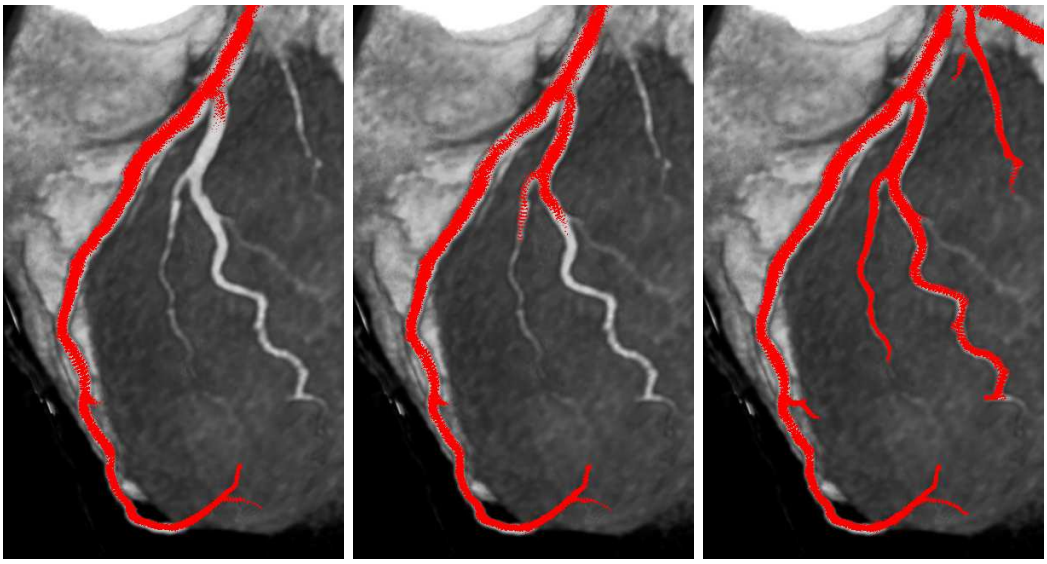


Figure 8.11: Effect of adaptive particle filtering and mean-shift detection. Left: classical SIR filter. Middle: our implementation of auxiliary particle filter with adaptive number of particles (AAPF). Right: AAPF with mean-shift detection of bifurcations and independent tracking of multiple branches. Successive particle populations (in red) are overlayed on a volume rendering of cardiac CTA data.

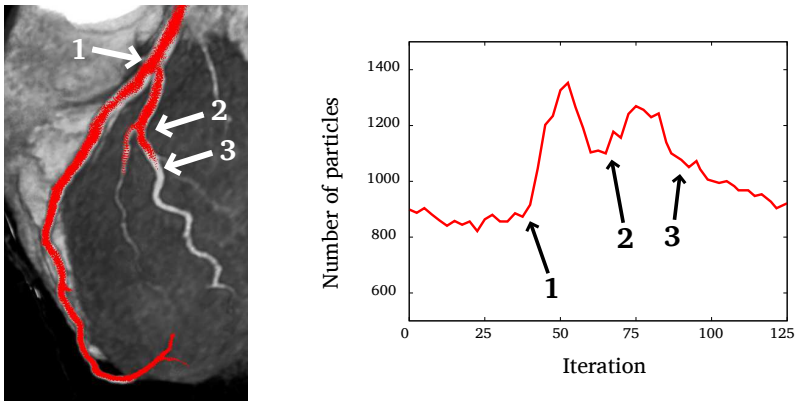


Figure 8.12: Adaptive number of particles and behavior at bifurcations. Left: successive populations of particles from an AAPF propagation (see also Fig. 8.11). Right: variation of the number N_t of allocated particles ($N^* = 500$) per iteration. One can notice an increase of N_t at bifurcations (arrows 1 and 2). This number progressively decreases (arrow 3) as secondary branches are lost because of sample impoverishment.

still lost after a while (Fig. 8.11, middle). In particular, one can note in Fig. 8.12 (right) that following a peak at bifurcations, the number of allocated particles tends to decrease as more and more particles are re-allocated to the dominant mode. Similarly to (Florin et al., 2005, 2006; Allen et al., 2008), we rely on an explicit detection scheme to extract each branch independently.

Figure 8.13: Mean-shift clustering. Given a set of samples $\{u^i \in \mathbb{R}^d\}_{i=1}^N$ (left), the mean-shift algorithm proceeds as a hill-climbing procedure on the estimated density $\hat{f}_h$ (right). Each sample u^i is moved iteratively towards a local maximum of the density (white circles). Final clustering is given by the assignment of the original samples to the detected maxima.

We emphasize that in Fig. 8.12, we selected a particularly explicit illustration of the behavior of our sample size adaptation scheme. In practice, N_t alone is not a robust indicator of the presence of a bifurcation, as other situations (noise, anomalies, lack of contrast) may also lead to increases of N_t. Where both Florin et al. (2005, 2006) and Allen et al. (2008) used K-means clustering, we preferred *mean-shift* mode detection (Fukunaga and Hostetler, 1975).

The use of *mean-shift* clustering for the detection of bifurcations is motivated from both a theoretical and a practical point of view. First, mean-shift procedures can be interpreted in terms of kernel density estimation, which integrates seamlessly with Monte-Carlo techniques. This allows us to preserve a homogeneous theoretical framework. Second, mean-shift clustering does not require an initialization and does not need the number of clusters to be specified. These are strong advantages over K-means. In our method, mean-shift serves two purposes. We use it for bifurcation detection and for the extraction of the final result (see Sec. 8.4.5).

8.4.3.1 Mean-Shift Clustering and Kernel Density Estimation

Mean-shift is an iterative, non-parametric procedure for unsupervised clustering and optimization. To the author's knowledge, its formulation was first given in (Fukunaga and Hostetler, 1975), but it really became popular with the work of (Comaniciu and Meer, 2002) on low-level computer vision tasks. It was since applied to a wide array of applications, from image filtering and segmentation, texture and feature-space analysis, to video tracking.

Mean-shift clustering relies on kernel density estimation (introduced in Sec. 6.2) for the analysis of a set of samples $\{u^i, w^i\}_{i=1}^N$. We consider the case of *weighted* samples $u^i \in \mathbb{R}^d$ with weights $w^i \in \mathbb{R}$ without loss of generality. The theoretical idea behind mean-shift clustering is the analysis of the density $\hat{f}_h(u)$ estimated from the samples:

$$\hat{f}_h(u) = \frac{1}{\sum_{i=1}^N w^i} \sum_{i=1}^N w^i K_H(u - u^i)$$

with K a d-variate kernel of bandwidth matrix H (see Sec. 6.2). The goal of mean-shift clustering is twofold:

- detect the local maxima $\{m^j\}_{j=1}^{n_m}$ of $\hat{f}_h(u)$ (*mode seeking*);
- assign each sample u^i to a mode m^j to obtain the final clusters.

Mean-shift techniques do not require the number n_m of modes, *i.e.*, clusters, to be specified.

Mean-shift procedures operate with a particular class of radially symmetric kernels of the form:

$$K(u) = c_k k(\|u\|^2)$$

with c_k a normalization constant ensuring integration to unity and k the so-called *profile* of K. In this work, we use the Epanechnikov kernel, whose profile is:

$$k_E(u) = \begin{cases} 1-u & 0 \leq u \leq 1 \\ 0 & u > 1 \end{cases} \tag{8.23}$$

Similar derivations can be obtained for the Gaussian kernel.

The motivation for the use of such kernels is to ease the estimation of density *gradients*. To this aim, we first introduce $g(u)$, the so-called *shadow* of $k(u)$, defined as:

$$g(u) = -k'(u)$$

Note that for the Epanechnikov profile (Eq. 8.23), the shadow $g_E(u)$ is uniform (d-dimensional unit sphere). The gradient of density at point u is given by (Comaniciu and Meer, 2002):

$$\hat{\nabla} f_h(u) = \frac{2c_k}{h^{d+2} \sum_{i=1}^N w^i} \left[\sum_{i=1}^N w^i g\left(\left\| \frac{u-u^i}{h} \right\|^2 \right) \right] \left[\frac{\sum_{i=1}^N w^i g(\|\frac{u-u^i}{h}\|^2) u^i}{\sum_{i=1}^N w^i g(\|\frac{u-u^i}{h}\|^2)} - u \right] \tag{8.24}$$

with h the bandwidth parameter. The second term is the so-called *mean-shift vector* $ms_h(u)$:

$$ms_h(u) = \frac{\sum_{i=1}^N w^i g(\|\frac{u-u^i}{h}\|^2) u^i}{\sum_{i=1}^N w^i g(\|\frac{u-u^i}{h}\|^2)} - u \tag{8.25}$$

The mean-shift *procedure* operates as an iterative gradient ascent for each sample u^j. Denoting by $\{v_k^j\}_{k=1,2,\ldots}$ the successive mean-shifted locations for sample u^j, the procedure is given by the following recursion:

$$v_0^j = u^j \tag{8.26}$$

$$v_{k+1}^j = ms_h(v_k^j) + v_k^j = \frac{\sum_{i=1}^N w^i g(\|\frac{v_k^j - u^i}{h}\|^2) u^i}{\sum_{i=1}^N w^i g(\|\frac{v_k^j - u^i}{h}\|^2)} \tag{8.27}$$

This procedure (illustrated in Fig. 8.13) is an *adaptive* gradient ascent scheme, as the stepping varies in function of the magnitude of the density gradient (Cheng, 1995; Comaniciu and Meer, 2002). Mean-shift moves are larger in areas of low density and smaller when approaching local maxima. Because of its adaptive nature, this algorithm is guaranteed to converge to stationary points of the estimated density (Cheng, 1995; Comaniciu and Meer, 2002). These stationary points are, in a vast majority of practical cases, local maxima (modes) of the density. Saddle points can be ruled out by checking the stability of the result, using small perturbations for instance. In practice, convergence is assessed when $\|v_{k+1}^j - v_k^j\| < \epsilon$, with ϵ a low distance threshold. We use $\epsilon = 0.05$mm in our implementation, for the clustering of the spatial positions $\{p_t^i\}$ of the particles.

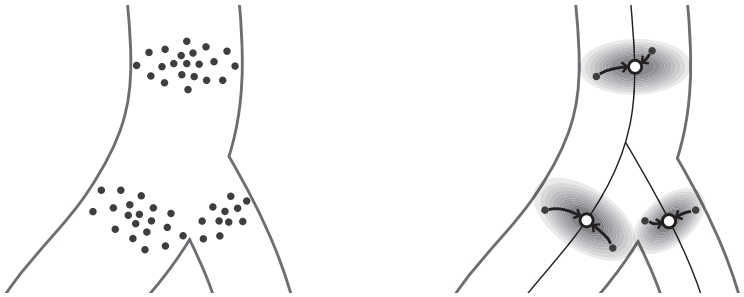


Figure 8.14: Mean-shift clustering for branching detection and centerline estimation. Left: at bifurcations, the population of particles (weighted spatial positions $\{p_t^i, \tilde{w}_t^i\}$) split into different clusters. Right: mean-shift clustering operates, for each initial position p_t^i, as a hill-climbing procedure towards local maxima of density of $\hat{f}_h(p_t|z_{1:t})$. The detected modes $\{\hat{m}_t^j\}_{t=0}^L$ (white circles) are used as an estimation of the centerline position.

8.4.3.2 Mean-Shift for Branching Detection

To detect coronary bifurcations, we apply mean-shift clustering on the particle *positions* with associated importance weights $\{p_t^i, \tilde{w}_t^i\}_{i=1}^{N_t}$ (see Fig. 8.14). This corresponds to a *marginalization* of our state space $x_t = (p_t, d_t, r_t)$. We recall that the Monte-Carlo importance estimation of the position posterior $p(p_t|z_{1:t})$ is given by:

$$\hat{p}(p_t|z_{1:t}) = \sum_{i=1}^{N} \tilde{w}_t^i \delta(p_t - p_t^i)$$

Our mean-shift procedure thus works on a continuous density estimated from $\{p_t^i, \tilde{w}_t^i\}_{i=1}^{N_t}$:

$$\hat{f}_h(p_t|z_{1:t}) = \sum_{i=1}^{N} \tilde{w}_t^i K_H(p_t - p_t^i)$$

This density is closely related to the Monte-Carlo approximation of the posterior density. A key difference is the regularization induced by the kernel.

Bandwidth Selection In this work, we use the Epanechnikov kernel for its simplicity and computational efficiency. As mentioned previously, the shadow (derivative) of this kernel is uniform, so that mean-shift computations reduce to simple weighted means. A key parameter, not yet discussed, is the bandwidth parameter h. As was mentioned in Sec. 6.2, adequate choices of bandwidth parameters are particularly intricate with non-Gaussian, multi-modal densities. From a statistical perspective, optimality results (Silverman, 1986; Fukunaga, 1990; Wand and Jones, 1995) are of low practical utility as they imply intractable quantities or rely on unrealistic assumptions. Detailed discussions on mean-shift bandwidth selection can be found in (Comaniciu et al., 2001; Comaniciu, 2003).

Fortunately, the injection of application-specific knowledge can help alleviate that issue. In our context, the data points $\{p_t^i\}$ correspond to spatial positions along vessels. Clustering is used to detect vascular bifurcations, *i.e.*, the apparition of new coronary branches. In our implementation, we simply set the bandwidth h_t to the Monte-Carlo

estimate of the expectation of the radius r_t at iteration t:

$$h_t = \hat{\mathbb{E}}[p(r_t|z_{1:t})] = \sum_{i=1}^{N_t} w_t^i r_t^i \tag{8.28}$$

The bandwidth value (radius of the Epanechnikov shadow) is adapted at each iteration. The Monte-Carlo estimate of the radius is mainly driven by the dominant branch, so that the strength of the regularization is adapted to the current (main) vessel of interest. We found this heuristic choice to provide a good tradeoff between false positives (spurious modes detected because of a too low value of h) and early detection of the bifurcations.

Computational Efficiency A counterpart to the robustness of mean-shift clustering is its relatively high computational cost (complexity of $O(N_t^2)$). Possible refinements include approximated variants (Vedaldi and Soatto, 2008), faster density estimations relying on binned estimators (Comaniciu and Meer, 2002) or fast Gauss transforms (Yang et al., 2003), schemes to accelerate the search of nearest neighbors, such as quadtrees and locality sensitive hashing (Georgescu et al., 2003) and kernel sampling techniques (Freedman and Kisilev, 2009).

For the sake of simplicity, we limited ourselves to a "naive" implementation of mean-shift. Nevertheless, we introduce some refinements to limit its computational impact on the overall tracking scheme. Our approach comes from a simple observation: bifurcations are relatively rare occurrences along a vessel. In a vast majority of cases, the density exhibits a single mode and full clustering is unnecessary. At each iteration of our filter, we propose to probe the density as follows:

- we first detect an initial mode m_t^1;
- we then sample the population of particles, looking for new modes $\neq m_t^1$. If at least one new mode is discovered, full mean-shift clustering is performed.

To detect the first mode m_t^1, we select the particle x_t^i of highest weight $\tilde{w}_t^i$ and apply the mean-shift procedure until convergence. Note that one could actually select any particle. Our scheme generally accelerates the procedure by selecting a particle likely to lie close to a local maximum of density. Second, we randomly select a small sub-population of particles $\{x_t^i\}_{i=1}^{N_p}$, with $N_p < N_t$. Test particles are sampled according to $\|p_t^i - m_t^1\|$, their distance to the first mode. Intuitively, the farthest a particle is from the first mode, the more likely it belongs to another (hypothetical) mode. The mean-shift procedure is run for each test particle sequentially. As soon as a test particle converges to a mode $m_t^2 \neq m_t^1$ ($\|m_t^2 - m_t^1\| > \epsilon$), we stop the probing process and perform full mean-shift clustering for all particles, as the existence of at least one secondary mode has been assessed. On our cardiac CTA data, actual bifurcations are present in less than 3% of the filter iterations on average. In other words, in 97% of the iterations, the N_p test particles will all converge to m_t^1. In our tests, fixing N_p as low as 5% of N_t was enough to robustly detect the presence of actual bifurcations. This means that in 97% of cases, our bifurcation detection scheme corresponds to applying mean-shift on only 5% of the particle population. This simple, empirical approach alleviates the burden of mean-shift clustering, keeping our overall algorithm at relatively good levels of computational efficiency (see Sec. 8.5).

When a bifurcation is assessed (with two, occasionally three modes detected), propagation is resumed independently for each cluster. Note that the re-population of each cluster is ensured naturally by the AAPF (line 9 in Fig. 8.10). Additionally, we update an inclusion mask with each newly extracted branch. Checking for inclusion in this mask avoids retracing the same branch twice.

8.4.4 Stopping Criterion

Similarly to our minimal path approach (see Sec. 7.6.2.2), we implemented a stopping criterion based on the likelihood ratios $\frac{p_v(y_t^i|r_t=r_t^i)}{p_{bg}(y_t^i)}$. We recall that a particle x_t^i whose likelihood ratio is inferior to 1 is considered as more likely to correspond to the background than to an actual vessel. A difference with the stopping criterion used in Sec. 7.6.2.2 is that we base our analysis on the entire population $\{x_t^i\}_{i=1}^{N_t}$ for increased robustness. If less than $F\%$ of the particles at iteration t have likelihood ratios lower than 1, the iteration is flagged. Propagation is stopped if more than half of the last S iterations are flagged. We empirically set $F = 25\%$ and $S = 20$ in our implementation. Results showed relatively low dependence on F in our experiments. The parameter S controls the tolerance to local anomalies. It allows the filter to locally loose the target vessel, wander and potentially re-capture it. Using $S = 20$ (approximate vessel length of $6mm$ for steps of $0.3mm$) proved to be a good compromise between premature stops and false positive detections (jumping into veins or ventricles, mainly).

8.4.5 Mean-Shift for Result Extraction

The successive generations of particles provide us with a Monte-Carlo approximation of the posterior distribution, from which we wish to extract a more compact, practical result representation. More precisely, we are interested in extracting the centerline curve and associated local radius. To that aim, a first possibility is to select, for each generation, the particle of highest weight(Florin et al., 2005, 2006). An alternative solution, corresponding to the true use of particle filters as an estimation tool (Doucet et al., 2001), is to extract the Monte-Carlo estimates of the expected centerline position $\hat{\mathbb{E}}[p_t|z_{1:t}]$ and radius value $\hat{\mathbb{E}}[r_t|z_{1:t}]$, as proposed in (Allen et al., 2008). The second approach is arguably more sound from a theoretical point of view, leading to more stable and smoother results. It is however sensitive to asymmetric tails in the distribution, which may yield some inaccuracies by shifting expectations away from local maxima[3]. In our implementation, we adopt an intermediate solution. We simply reuse the successive mean-shift modes as an estimation of the centerline curve[4] (see Fig. 8.14). From a theoretical point of view, these modes correspond to the local maxima of the estimated density. Kernel regularization ensures smoother results than simply selecting the best particle, while the estimation is more local, generally slightly more accurate than the corresponding Monte-Carlo estimate. Sample results are illustrated in Sec. 8.5.

The radius estimation associated to a mean-shift mode m_t is given by:

$$r_{m_t} = \sum_{i=1}^{N_t} \hat{w}_t^i K_E(\frac{m_t - p_t^i}{h_t}) r_t^i \tag{8.29}$$

with K_E the Epanechnikov kernel, with the same bandwidth h_t used for the mean-shift procedure. This estimate is again more local and, in our tests, slightly more accurate than the Monte-Carlo one.

Finally, one can notice that our result representation (chain $\{m_t, r_{m_t}\}$ of successive

3. Since each branch is tracked independently, we expect the position and radius posterior distributions to be generally mono-modal, but not necessarily symmetric.

4. Even if no bifurcations were detected, we extracted the position of the mode.

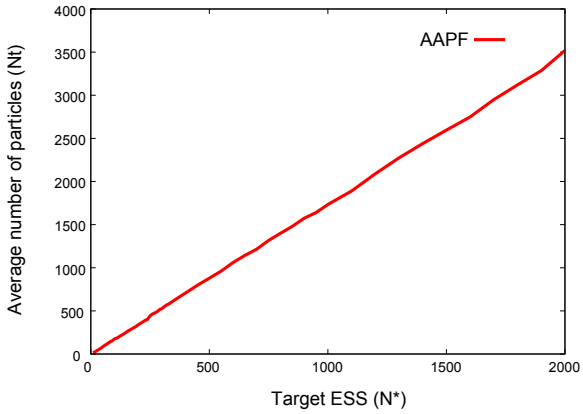


Figure 8.15: Average number of particles $\bar{N}_t$ as a function of the target ESS N^*, for CTA coronary artery tracking. Illustrative result for one test dataset, for which $\bar{N}_t \simeq 1.75N^*$.

modes and radiuses) is not a realization of our model [5], but the result of numerical integrations over the realizations considered by the filter.

8.5 Experiments and Evaluation

8.5.1 Experiments

To evaluate the behavior of our AAPF algorithm, we first conducted a series of experiments on a single dataset with good image quality, but with a particularly complex left coronary tree, with numerous branches of various sizes and curvatures (see Fig. 8.17).

The evaluation criteria, such as the dice coefficient of overlap (OV) and the centerline distance (AI) to the ground-truth delineation, were detailed in Sec. 7.7.2.1.

8.5.1.1 Target Effective Sample Size N^* and Number of Allocated Particles

A key parameter of our AAPF algorithm is the target effective sample size N^* used for the dynamic adaptation of the number N_t of particles allocated at each iteration (see Sec. 8.3.2). In our tests, we observed variations of N_t between approximately $1.5N^*$ up to $5N^*$ at distal ends of thin, noisy coronaries. Fig. 8.15 depicts the average number of particles $\bar{N}_t$ as a function of N^* for our selected dataset. The relation is almost exactly linear, with $\bar{N}_t \simeq 1.75N^*$. Tests on different datasets led to similar results, with slightly different linear slope coefficients, depending mainly on the image quality. Expectedly, lower image qualities tend to induce larger coefficients.

8.5.1.2 Comparison of SIR, APF and AAPF

The performance of our AAPF algorithm was compared with classical sampling-importance-resampling (SIR) and auxiliary particle filter (APF) implementations. All three algorithms employ explicit detection of bifurcations by mean-shift clustering (Sec. 8.4.3). For the sake

5. Monte-Carlo estimates also would not guarantee that the spacing of the extracted points and their orientations satisfy exactly the properties of our geometric model.

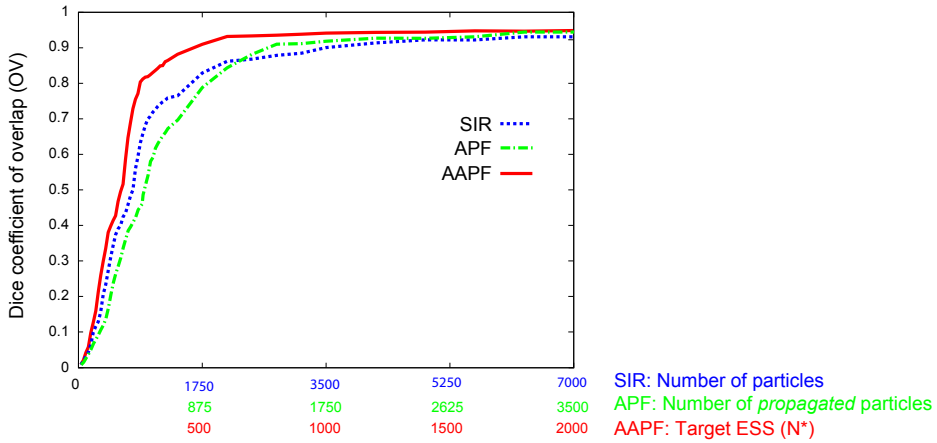


Figure 8.16: Average Dice coefficient of overlap with ground truth (OV) for SIR, APF and AAPF algorithms. Results averaged on 50 independent runs on a single cardiac CTA data set. Each algorithm uses a number of particles corresponding to an equivalent level of computational performance (see text).

of fairness, the different algorithms are compared for equivalent levels of *computational* cost. AAPF propagates a varying number of particles N_t at each iteration, determined with respect to the target effective sample size N^*. One iteration of the AAPF, consisting of one simulation pass and one actual propagation, is equivalent to the evaluation of $2\bar{N}_t$ particles, with $\bar{N}_t$ the average number of particles induced by N^*. The number of particles for each algorithm is thus set as follows:

- AAPF uses a target effective sample size N^*;
- APF uses $\bar{N}_t$ particles for simulation (see Sec. 8.3.1) and another $N_{APF} = \bar{N}_t$ for the actual propagation;
- SIR uses $N_{SIR} = 2\bar{N}_t$ particles.

The number of particles propagated by the APF and AAPF (on average) is thus twice lower than for SIR.

Fig. 8.16 shows the average Dice coefficient of overlap with the ground truth delineation (OV) as a function of the number of particles for the three algorithms. As expected, increasing the number of particles improves the performance of all three algorithms. Interestingly, up to $N_{SIR} \simeq 2000$ particles, the performance of the SIR algorithm is superior to that of the computationally equivalent APF. Auxiliary sampling indeed improves the allocation of the samples, but for relatively low numbers of particles, it is actually advisable to sample twice as many SIR particles for the same computational cost. For $N_{SIR} > 2000$, the APF slightly outperforms classical SIR. The benefit from the dynamic adaptation of the sample size is obvious. The AAPF algorithm clearly outperforms SIR and APF, even for relatively low numbers of particles. We attribute this result to a better capture of secondary branches by the AAPF, even for relatively low values of N^* (see Sec. 8.4.3.2). The same observation was drawn during our quantitative evaluation on our entire database (see Sec. 8.5.2). All three algorithms reach some kind of performance plateau ($OV \simeq 0.93$ for this test dataset). As illustrated in Fig. 8.17 (middle), typical runs consistently fail at extracting one thin, secondary branch, independently of the number of particles employed. We believe such issues to be linked to our model, as discussed in Sec. 8.5.2.

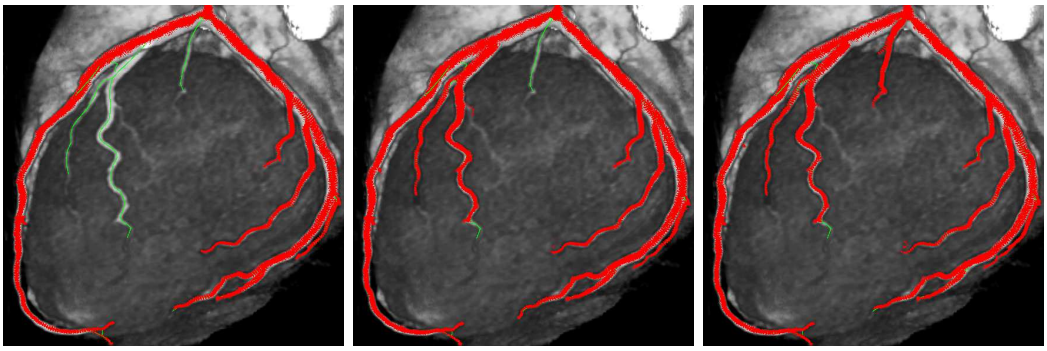


Figure 8.17: Different runs of the AAPF on the same dataset ($N^* = 1000$). Left: an isolated lower quality result ($OV = 0.83$) with several missed branches, obtained in $\sim 4\%$ of the runs. Middle: a typical run ($OV = 0.93$). Right: a high quality run ($OV = 0.97$), obtained in $\sim 30\%$ of the runs. Ground truth centerline delineation is depicted in green.

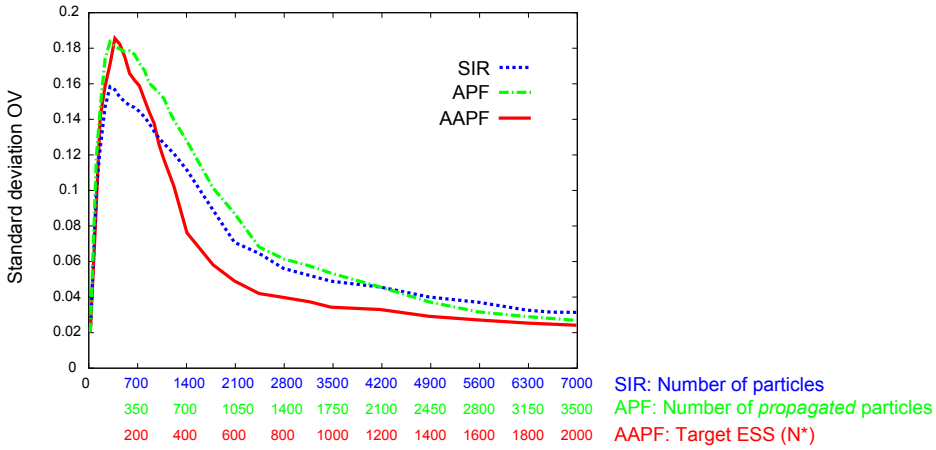


Figure 8.18: Standard deviation of the Dice coefficient of overlap with the ground truth (OV) for SIR, APF and AAPF algorithms. Results obtained for 50 independent runs on a single cardiac CTA data set. Each algorithm uses a number of particles corresponding to an equivalent level of computational performance (see text).

An important practical issue with stochastic approaches is the consistency of the extracted result over runs. As illustrated in Fig. 8.17, different runs may yield different results. Limited variations of the estimated centerline are acceptable. On the other hand, the fact that the algorithm may sometimes miss a branch entirely is much more problematic. In Fig. 8.18, we give the estimated standard deviation of the overlap OV as a function of the number of particles. For all three algorithm, a peak of result variance is obtained for low numbers of particles. These peaks corresponds to the points where the filters start having sufficient statistical power, in most cases, to extract significant portions of the coronary tree (see Fig. 8.16). We emphasize that a standard deviation of 0.16, for a Dice coefficient, is considerable. This corresponds to the loss or gain of several secondary branches. As expected, the variance decreases for larger numbers of particles. The SIR appears to have the slowest decrease rate, but APF and AAPF have slightly higher initial

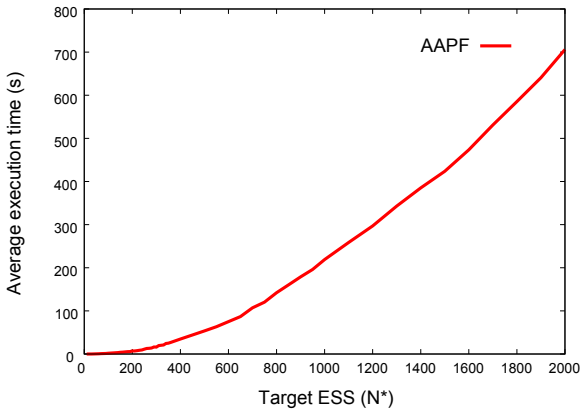


Figure 8.19: Average computational time as a function of N^*. Result averaged on 50 independent runs.

variances. For reasonable values of N^*, the AAPF exhibits a lower variance than both other algorithms. In practice, for $N^* = 1000$, the results obtained by the AAPF are fairly stable, with very little difference from run to run. The variance is however non-null and the algorithm may still miss, or gain (see Fig. 8.17, right) a branch with respect to a typical run. These variations concern mainly thin secondary branches, which account relatively small fractions of the overall coronary trees. Main coronary branches are most generally extracted in a very consistent manner, even for relatively low values of N^* [6].

8.5.1.3 Algorithmic Complexity and Execution Time

The AAPF involves only $O(N_t)$ operations and we showed in Sec. 8.5.1.1 that the relation between the average number of particle $\bar{N}_t$ and N^* is empirically linear, with a multiplicative constant depending on the data at hand. Our implementation of mean-shift clustering is however of quadratic complexity $O((N_t)^2)$. Fortunately, the algorithmic refinements discussed in Sec. 8.4.3.2 alleviate its impact on the overall execution time. As illustrated in Fig. 8.19, the average execution time increases in a non-linear, but less than quadratic function of N^*. Please note that for low values of N^* ($N^* < 400$), the execution time is understandably very low as the filter is generally unable to explore the entire vascular tree.

Despite our refinements, the computational overhead induced by the use of mean-shift clustering remains relatively high. Bifurcation detection accounts for approximately 20% of the overall execution time. We emphasize however that our approach remains reasonably fast, thanks to the computational efficiency of the underlying components (`MFlux` feature in particular). For $N^* = 1000$, a full coronary tree is generally extracted within 4 minutes (2.16GHz `Intel Core Duo Intel` processor). Given the numerous possibilities for improvement (massive parallelization and optimization of mean-shift clustering), we believe that the performance of our proof-of-concept implementation is particularly promising.

6. In our tests, a value of $N^* = 500$ seemed sufficient for the reliable extraction of the main coronary arteries.

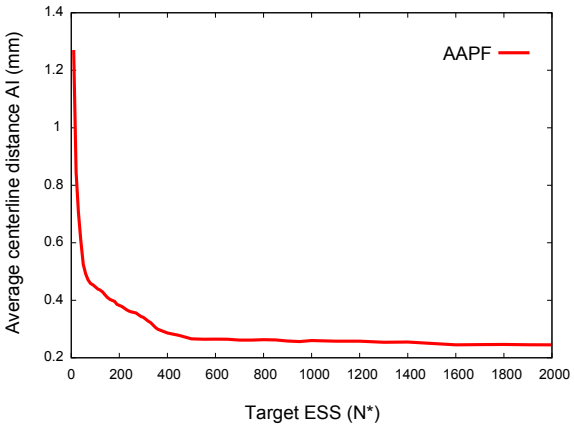


Figure 8.20: Average centerline distance (AI) to the ground truth as a function of N^*. Result averaged on 50 independent runs.

8.5.1.4 Centerline Accuracy

As depicted in Fig. 8.20, the average centerline error (AI) decreases relatively quickly with increasing N^*. For $N^* \simeq 600$, AI reaches $\simeq 0.26mm$, which is a satisfyingly subvoxelic precision (the intra-slice resolution of the test dataset is $0.33mm$). Increasing N^* further brings only marginal improvement of the centerline accuracy ($0.248mm$ for $N^* = 2000$).

Measure	SIR ($N_{SIR} = 2\bar{N}_t$)	APF ($N_{APF} = \bar{N}_t$)	AAPF ($N^* = 1000$)
FN	21.6%	19.3%	15.4%
FP	12.5%	12.1%	12.3%
OV	83%	84.3%	86.2%
Std. dev. OV	4.7%	3.9%	3.2%
OT	92.1%	92.4%	92.5%
Std. dev. OT	1.7%	1.7%	1.5%
AI	0.26mm ($\pm$0.22)	0.26mm ($\pm$0.21)	0.25mm ($\pm$0.21)
AR	0.21mm ($\pm$0.19)	0.20mm ($\pm$0.18)	0.20mm ($\pm$0.18)

Table 8.2: Particle filtering extraction: internal quantitative validation on 51 datasets. Results averaged on 20 runs for each dataset. See text for details.

8.5.2 Evaluation

Our algorithm was quantitatively evaluated on our database, similarly to our minimal path approach in Chapter 7. We used 10 randomly selected datasets for training (*i.e.*, learning the likelihood and prior distributions) and the 51 remaining ones for testing. We evaluated the extraction of both left and right entire coronary trees. Sample results are given in Fig. 8.21 and 8.22.

The quantitative evaluation criteria are identical to those used in Sec. 7.7:

- FN: average percentage of false negatives, *i.e.*, portion of the ground truth delineation not successfully extracted by the algorithm;
- FP: average percentage of false positives, *i.e.*, portion of the extracted result not present in the ground truth delineation;
- OV: average overlap ratio (Dice coefficient);
- OT: average overlap ratio for clinically relevant vessels (diameter $\geq$ 1.5mm);
- AI: average centerline error (distance to the ground truth delineation, for true positive portions of the result, standard deviation in parentheses);
- AR: average radius estimation error (for true positive portions of the result, standard deviation in parentheses);

The reader may refer to Sec. 7.7.2.1 for more in-depth discussions of these criteria. In addition, we also report the average standard deviation of OV and OT *among runs* on the same dataset, to evaluate the practical consistency of the method.

Table 8.2 summarizes the quantitative results for the SIR, APF and AAPF algorithms. These results were obtained for $N^* = 1000$ for the AAPF, ensuring sufficient statistical power and low variance of the results (see Sec. 8.5). For each dataset, we first ran the AAPF, computed the average number $\bar{N}_t$ of particles used, and ran the SIR algorithm with $N_{SIR} = 2\bar{N}_t$ particles and the APF with $N_{APF} = \bar{N}_t$.

Comparison of AAPF, APF and SIR The quantitative results confirm the experiments of Sec. 8.5, with AAPF outperforming APF and SIR in terms of robustness (overlap criteria) and leading to a marginal improvement of centerline accuracy. The AAPF approach yields notably an appreciable decrease of the rate of false negatives (missed segments) and inter-run variance, confirming that our adaptive method is able to explore the arterial tree in a more consistent and exhaustive manner.

Overlap measures (OV and OT) From a quantitative point, the AAPF approach fares actually slightly better than our minimal path approach (Sec. 7.7.2), with an average OV percentage of 86.2 (against 85). This result is mostly credited to a lower rate of false positives (12.3% versus 16.5%). In practice, particle filtering seems less prone to leaks into veins and ventricles than the minimal path method. We believe this comes mainly from the use of the directional prior. Additionally, particle filters are much less susceptible to shortcut issues, as they rely on direct, iterative tracking. On the other hand, the rate of false negatives is slightly higher than for minimal paths (15.4% versus 13.5%). The AAPF algorithm seems to miss more secondary asymmetric bifurcations than the minimal path approach. Missing or gaining such branches also explains the average inter-run standard deviation of OV (3.2%). This issue is discussed more in-depth in a subsequent paragraph. We highlight however the good consistency of the result for main coronary branches (diameter $> 1.5mm$), with an overlap Dice coefficient OT of 92.5% and with low inter-run deviation (1.5%)

Centerline and radius accuracy From a qualitative point of view, results extracted thanks to the modes of the mean-shift procedure are satisfyingly regular, following closely the ground truth delineations even in curvy portions of the coronaries. The algorithm can also be seen as adapting well to radius variations (see the right columns of Fig. 8.21 and 8.22). Compared to the minimal path approach of Chapter 7, the particle filtering method does not depend on a discretization of the search space. Quantitatively, this yields an improvement of the centerline error (AI), from 0.31 to 0.25mm on average. This error is satisfyingly sub-voxelic, confirming the results of Sec. 8.5. Increasing the parameter N^* over 1000 only resulted in marginal improvements in our tests, suggesting that this level of accuracy is probably more dependent on the low-level `MFlux` feature, and notably on the assumption of circular cross-sections, than on the extraction scheme. The average error of radius estimations (AR) obtained with AAPF is on par with that of the minimal path method.

On asymmetric bifurcations One of the main issues observed with our AAPF algorithm is that, regardless of the number of particles employed, the filter often misses some of the smallest coronary branches, arising from asymmetric bifurcations. These secondary arteries generally branch from much bigger vessels, at relatively high angles. Such occurrences are in fact poorly predicted by our direction and scale priors, which correspond to relatively smooth, limited angular and scale variations along the vessel. One can note, in the left columns of Fig. 8.21 and 8.22), that some of the secondary branches explored by the filter are originally discovered by a small number of particles, predicted in the tails of our transition priors. For most of these branches, adequate statistical power, *i.e.*, a sufficiently high number of sampled particles, is enough to guarantee their consistent detection. This is not the case, however, for those with particularly high scale variations (*e.g.* a very small coronary branching off a big one) *and* high angles of bifurcation. We believe that this is not a problem of sample impoverishment, but a more fundamental issue of our model. The algorithm has difficulties capturing some branches because it rarely samples them at all. Because the prior performance is in question, Markov Chain Monte Carlo (MCMC) rejuvenation, proposed in (Allen et al., 2008) for the re-population of bifurcations, would not solve the theoretical issue at hand. MCMC steps cope with particle impoverishment, but still rely on the same transition prior as the main filtering process. Practical improvement in our case can empirically be obtained by artificially

relaxing the transition prior, allowing broader angular and scale variations, but at the risk of making the exploration of the search space less efficient and more prone to false positives. Alternatively, one could probably devise a reseeding procedure to recover missed branches by exploring the immediate surroundings of the initial result. The use of a corner detector, as proposed in (Fridman, 2004), might help focus the algorithm's attention to potential branchings. From a theoretical point of view, an interesting lead would be the use of *model-switching* approaches (Isard and Blake, 1998a) that would allow us to take into account an explicit model for asymmetric bifurcations. The general idea would be to design a model-switching process enabling, at each iteration the prediction of a certain quantity of *branching* vessels, following a different transition prior (lower scales and high branching angle). Adequate switching and second transition priors could potentially be learned similarly to our current model.

8.6 Discussion and Perspectives

In this Chapter, we have presented our particle filtering approach to the extraction of full coronary trees from cardiac CTA data. This method relies on the stochastic propagation of a population of samples along the coronaries, following a centerline-based tracking design. A key difference with classical deterministic tracking schemes is that particle filters track the entire posterior distribution of the vessels. They consider numerous realizations in parallel from which they approximate the distributions and estimates of interest.

Our method was motivated by the seminal idea of (Florin et al., 2005, 2006) on particle filters for vascular tracking and by the subsequent works it inspired (Schaap et al., 2007a,b; Allen et al., 2008). Our approach relies on our full Bayesian model, including radius and direction priors learned from our ground truth database. We introduced an adaptive sampling scheme, which we referred to as Adaptive Auxiliary Particle Filtering (AAPF). AAPF combines classical auxiliary filtering (APF) with a dynamic adaptation of the number of samples. In our application, we showed that, compared to classical particle filters, this algorithm improves the behavior at bifurcations and lowers the stochastic inter-run variance of the result. We also discussed the use of mean-shift clustering for the explicit detection of bifurcations and for the extraction of the final result. Our quantitative validation demonstrated the robustness of this approach. Compared to the minimal path technique of Chapter 7, our particle filtering method yields qualitatively smoother, quantitatively more accurate results. On the other hand, it is less versatile in terms of possible workflows, as it is not suited for two-point interactive segmentation. It suffers from a slightly higher rate of false negatives (*i.e.*, missed vessel parts), and is computationally more intensive in its current implementation.

We believe that the results obtained by our prototype are particularly promising. From a theoretical point of view, we previously mentioned the perspective of using a model-switching procedure (Isard and Blake, 1998a) to improve the exploration of small, secondary branches. Another theoretical lead for improvement would be the use of *smoothing* estimation schemes (Kitagawa, 1996; Hürzeler and Künsch, 1998; Doucet et al., 2001), *i.e.*, estimations taking into account *future* observations (in opposition to *filtering* techniques, which rely solely on past observations). The idea of using future observations is particularly attractive in our applicative context, where the observation source (the image) is static and fully accessible. Smoothing techniques are considerably more computationally demanding than their filtering counterparts, but recent developments have shown promising optimization techniques (Klaas et al., 2006). An alternative, which could integrate

directly within our framework, would be the use of dynamic programming methods to extract the MAP sequence from successive particle generations (Godsill et al., 2001). Such a refinement would also induce considerable computational overhead.

Finally, we emphasize that our current implementation could be optimized in numerous ways. One could first accelerate mean-shift clustering thanks to some of the various existing refinements (Vedaldi and Soatto, 2008; Comaniciu and Meer, 2002; Yang et al., 2003; Georgescu et al., 2003; Freedman and Kisilev, 2009). A second, highly promising lead for optimization is the parallelization of the algorithm. In fact, most operations involved in our approach can be straightforwardly parallelized. This property is a strong computational advantage over the minimal path approach of Chapter 7. In our tests, we obtained a 35% performance gain by simply parallelizing the execution of the filter on two processor cores (`OpenMP` directives on a `Intel Core Duo Intel` processor). We believe that a massively parallel GPU implementation could reach very high levels of computational efficiency.

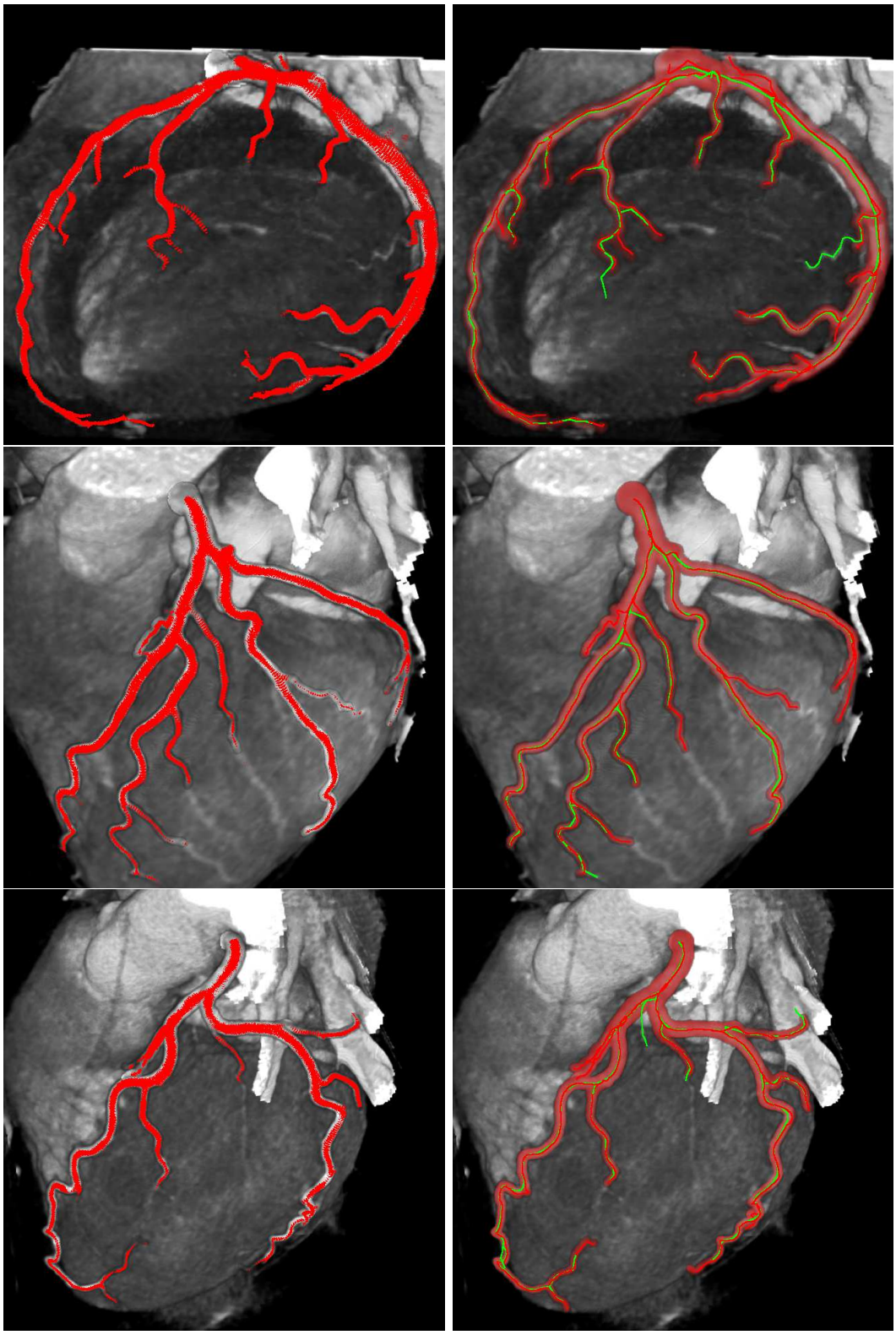


Figure 8.21: Tracking with particle filtering: result samples for left coronary trees. Left: successive generations of particles. Right: the extracted centerline is given in red. The associated mask (light red) is obtained from the corresponding radius estimation. The ground truth centerline is depicted as a green curve.

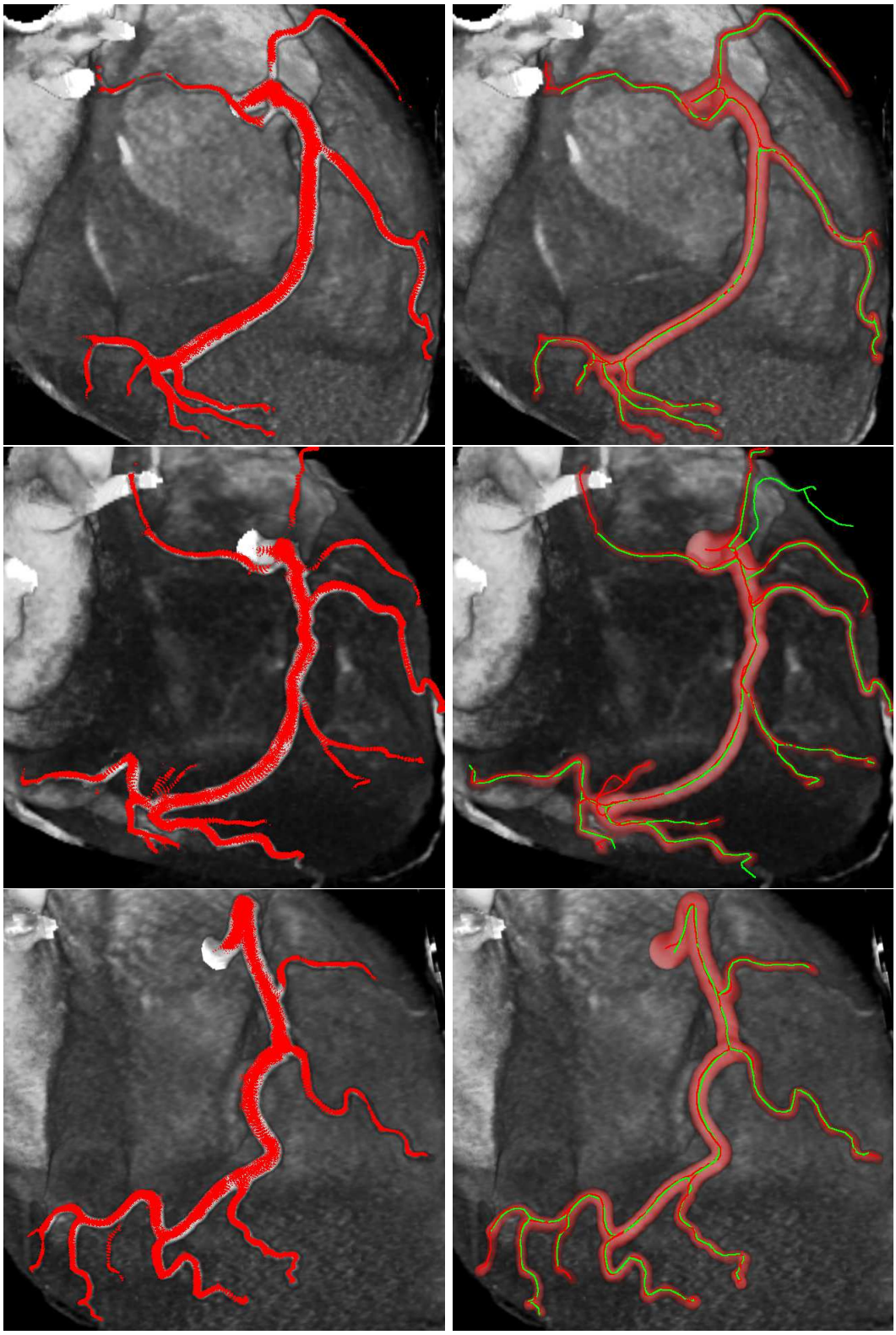


Figure 8.22: Tracking with particle filtering: result samples for left coronary trees. Left: successive generations of particles. Right: the extracted centerline is given in red. The associated mask (light red) is obtained from the corresponding radius estimation. The ground truth centerline is depicted as a green curve.

Chapter 9

Conclusion and Perspectives

9.1 Contributions

In this work, we discussed the key design issues of vascular segmentation methods along three axes: models, features and extraction schemes. We discussed what type of information can be modeled and incorporated as prior knowledge, how this information can be evaluated on the image and how model- and image-based data are eventually combined and optimized to perform the segmentation in practice. From an applicative point of view, our work focused on a particularly challenging and medically important problem, the segmentation of coronary arteries from cardiac CTA data.

A first contribution of this thesis is an extensive bibliographic review of state-of-the-art methods for 3D vascular segmentation. Our three-axis analysis differs from existing reviews which commonly rely on a linear categorization of works with respect to their mathematical frameworks. More than a direct comparison of existing techniques, our focus was on highlighting general trends in the literature and discussing their theoretical and practical properties. We believe that this high-level point of view contributes to a more general understanding of the field.

In Chapter 4, we introduced a discrete medial-based model for the geometric description of tubular structures. Its specific parameterization scheme allows us to compactly encode centerline, radius and directional information. We used that geometric representation to model coronary arteries in 3D CTA data and additionally studied assumptions on the appearance of coronary arteries in such data. To evaluate our geometric and appearance model on images, we developed a new multiscale oriented medialness feature (Chapter 5). Its key properties are its simplicity, its computational efficiency and its discriminative power. In particular, we showed that the introduction of a non-linear component in our seminal flux-based formulation considerably decreases false positive responses by enforcing symmetry assumptions.

In Chapter 6, we presented our recursive Bayesian model, built upon our geometric model and our vessel-dedicated feature. Its *a posteriori* formulation integrates model-based prior knowledge on radius and direction variations along coronary arteries. Its likelihood observation model combines statistical knowledge of typical feature responses on CTA coronaries *and* in the background. One key contribution of our model is its non-parametric formulation. All its components are learned from a database of manual delineations, avoiding the use of empirical parametric models. We showed that kernel density estimation integrates seamlessly in this Bayesian context and enables the capture of subtle information, such as dependencies of geometric and appearance variations on the

vessel radius.

From this Bayesian formulation, we derived two extraction strategies, reflecting alternative design choices. The first one, detailed in Chapter 7, relies on discrete graph-based minimal path optimization. It builds on recent developments of multiscale, 4D (centerline+radius) schemes. A first focus of this work was the design of a suitable cumulative cost metric, which we empirically derived as a simplification of our Bayesian model. In this regard, our approach avoids parameterization issues common to most related techniques. Second, particular attention was given to the computational efficiency of this method. We proposed different refinements to limit the amount of scale-space exploration and discussed key implementation details. In both interactive and automatic workflows, we demonstrated the robustness and flexibility of this extraction strategy through qualitative and quantitative evaluations. The two-point interactive version of the algorithm reaches very high robustness, with overlap Dice coefficients with ground-truth over 97%. The results obtained by the fully automatic version of the algorithm are also very promising, with overlap coefficients over 85% (internal validation with fully delineated trees) and 91% (Rotterdam evaluation (Schaap et al., 2009a) on four main vessels per datasets). Overall, our quantitative evaluation showed that both versions of our minimal path approach rank well in comparison with other state-of-the-art techniques. We believe that our method proposes an attractive balance between robustness, accuracy and computational efficiency.

Our second extraction strategy (Chapter 8) employs particle filtering to iteratively track coronary arteries. Compared to the minimal path approach, this stochastic, multi-hypothesis scheme does not require a discretization of the search space, yielding an improvement of the accuracy of the results. Where the minimal path approach employs an empirical simplification of our Bayesian mode, our particle filtering strategy effectively estimates the posterior distribution of our full model. On the other hand, it is restricted to a direct tracking workflow as it does not allow for two-point, constrained interactive processing. It is also computationally more intensive, although it could be massively parallelized. Our implementation improves on classical particle filters through its fully adaptive sampling scheme and the use of mean-shift clustering for branching detection and result extraction. Through a series of experiments, we showed that the combination of these techniques increases tracking robustness and reduces inter-run stochastic variance. Quantitative results for our particle filtering approach are comparable to those of our minimal path strategy, with a slightly lower false positive rate (less leakage issues) but a higher false negative rate (missed branches).

Throughout this work, we demonstrated the combination of careful design choices with various algorithmic refinements to obtain robust, yet computationally efficient segmentation methods. As detailed in Appendix C, all the tools developed during this thesis were implemented with robustness, computational efficiency and re-usability as central concerns. Another strong focus of attention was systematic, qualitative and quantitative evaluation of our approaches on a large array of clinical data. We notably developed a series of generic tools to automate such testing procedures. Our validation framework is currently being re-used in various projects within Siemens Corporate Research.

9.2 Perspectives

In the respective discussions of each chapter, we mentioned various theoretical and practical leads for the improvement of our approaches. First, our `MFlux` feature could be extended to take into account local narrowing and widening, following for instance the ideas of the Core technique from (Fridman, 2004). A possible extension of our Bayesian model would be the addition of a specific model for bifurcations. Our current model corresponds to single branches and does not well account for the brutal changes in orientation and radius which happen at asymmetric bifurcations, where a small secondary artery branches off a main coronary with a high bifurcation angle. The design of such a bifurcation model could help improve the capture of small branches, especially by the particle filtering strategy.

For the minimal path approach, one important point of possible improvement would be the reduction of false negatives (premature stopping) in cases of severe stenoses and noisy, faint vessels. The robustness to premature stopping is to be balanced with leakage issues (false positives) into blood-filled ventricles and vein network which can affect our fully automatic algorithm. To this end, a possibility would be to refine our stopping criterion and/or to integrate our minimal path propagation scheme into a higher-level framework. For instance, one could imagine trial-and-error re-seeding and pruning strategies, where the overall coherence of the extracted result would be controlled by a global, graph-based algorithm such as the reconnection and cleaning schemes employed in (Bauer and Bischof, 2008a; Szymczak et al., 2005). Potential computational optimizations include the implementation of a more efficient heap management scheme such as (Yatziv et al., 2006) and the use of bi-directional and A^* heuristic schemes for the two-point interactive workflow. Also, it will be interesting to explore possible workflow refinements for the interactive variant of the algorithm, such as allowing the definition of multiple intermediate seed points.

One of the issues affecting the particle filtering approach is a relatively high false positive rate, due in particular to its inability to capture some of the smallest secondary branches. As discussed previously, we explain this observation by the inadequacy of our Bayesian model, more specifically of the transition priors it embeds, at asymmetric bifurcations. The derivation of a bifurcation specific model could help alleviate that issue. Such a secondary model could be integrated within the particle filtering scheme thanks to a model-switching approach, as proposed in (Isard and Blake, 1998a). From a computational point of view, we highlight the high potential of a massively parallel GPU implementation.

Both extraction strategies could also benefit from subsequent post-processing steps. The accuracy of the extracted results by our techniques is currently of the order of the data resolution. This result is in accordance with our primary goal, which was to focus on the robustness of the delineation before its accuracy. We believe the accuracy is mainly limited by the relatively low-level feature employed, which is notably constrained by the assumption of circular cross-sections. It could potentially be improved by refining locally the initial centerline position. One possibility to achieve that goal would be by the extraction of an accurate lumen segmentation, as discussed thereafter.

From a general point of view, our work focused on a preliminary but crucial task, which is the robust delineation of the vessels of interest. The next logical step is thus the integration of our techniques into complete clinical workflows. As illustrated throughout this document, our extraction results can be directly exploited to help screening routines,

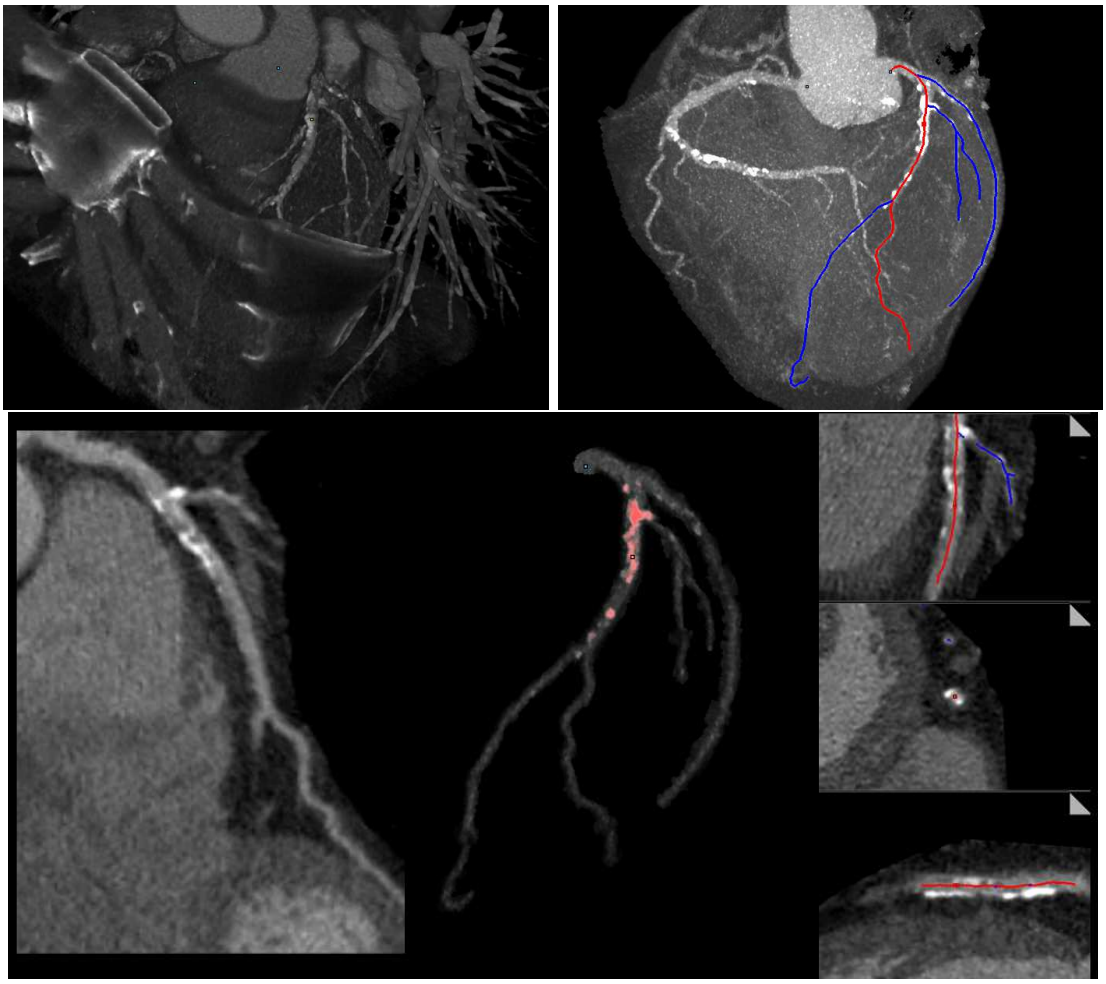


Figure 9.1: Integration of centerline extractions in screening workflows. Top left: volume rendering of the original CTA dataset. Top right: MIP view obtained after the segmentation of the heart surface, of the aorta and of the heart chambers (CathLab view). Areas outside the heart and inside the heart chambers are masked to facilitate the visualization of the coronary arteries. Extracted centerlines are overlayed, with the red branch being selected. Bottom row, from left to right: CPR view of the selected branch, local MIP view, orthogonal MPR view. In the MIP view, calcified areas are highlighted (threshold > 500H.U.).

notably through the automatic generation of CPR and local MIP views (see Fig 9.1). The rough volume segmentations provided by our algorithms may also be sufficient for calcium scoring. Other applications such as stenosis quantification potentially require a more accurate segmentation of the vessel lumen, which is fortunately facilitated by the pre-extraction of the centerlines and local radius estimations. In Sec. 3.3.6 of our bibliographic study, we discussed techniques which could be used to obtain such a refined lumen segmentation. In Fig. 9.2, we illustrate preliminary experiments with the mean graph cycle optimization from (Gulsun and Tek, 2006, 2008b,a). This technique is used to segment 2D contours in cross-sectional planes given by our extracted centerlines. The 3D mesh is constructed from the triangularization of successive 2D contours. A promising,

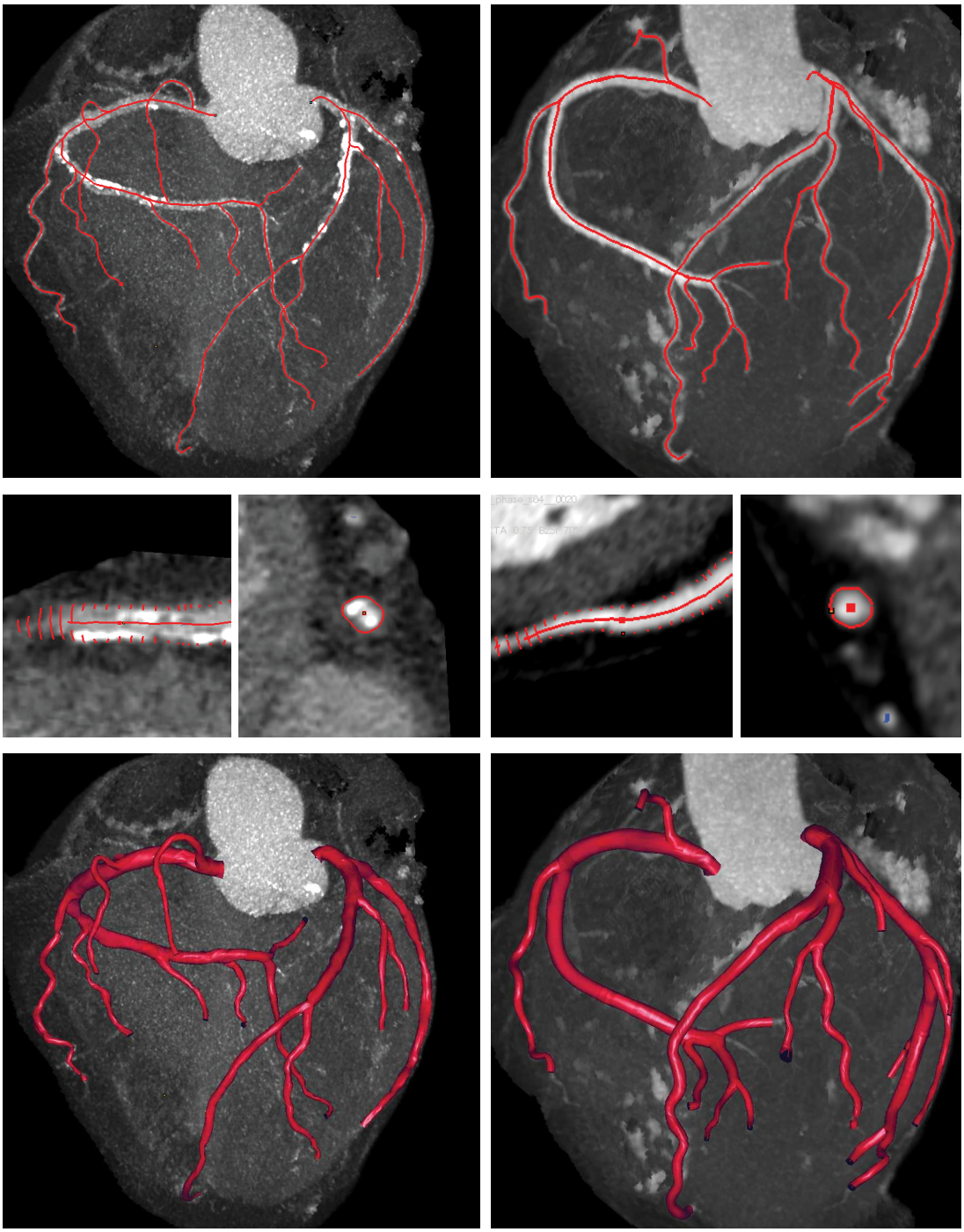


Figure 9.2: Automatic generation of a 3D surface model from extracted centerlines. We re-used the method of (Gulsun and Tek, 2006, 2008b,a) and illustrate the approach for two datasets, one per column. Top row: MIP CathLab view (see Fig. 9.1) with overlayed centerlines. Middle row: 2D cross-sectional contours are segmented in planes orthogonal to the centerlines. Bottom row: the 2D contours are triangularized to obtain a 3D mesh.

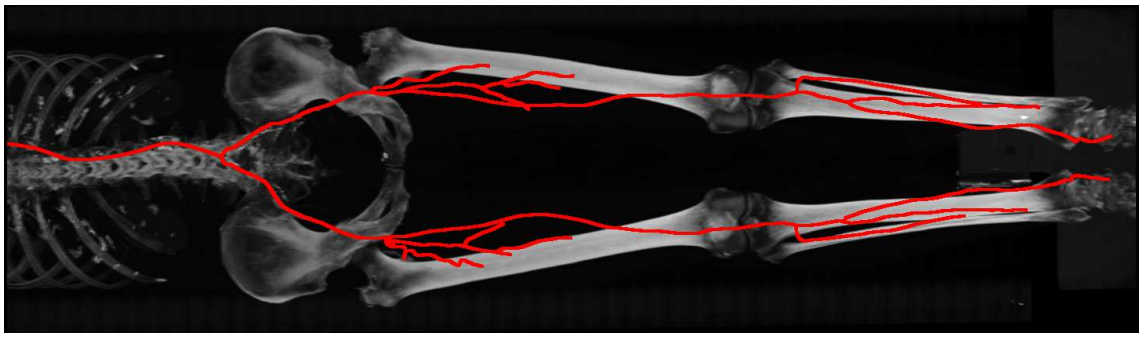


Figure 9.3: Preliminary result on peripheral CT data. Segmentation of the descending aorta and iliac arteries using our minimal path approach (1 seed in the aorta, 15 distal seeds in the iliac arteries).

fully 3D alternative can be found in the recent graph-cut-based approach of (Schaap et al., 2009b), which was successfully applied to CTA coronaries. Besides quantification tasks, an accurate contour segmentation could help refine the centerline position locally and thus improve the accuracy of our results.

Finally, we believe our methodological framework to be general enough to be easily adaptable to vascular applications other than CTA coronary arteries. The main bottleneck to the application of our learning-based approach is the need for manual delineations to be available. In the absence of a ground truth database, one can still resort to empirical, parametric models. For instance, we illustrate in Fig. 9.3 the application of our interactive minimal path approach to peripheral CT data. To obtain that result, we simply increased the maximum possible radius to 5cm and used a Normal distribution $\mathcal{N}(.|r_{t-1}, \sigma^2)$ of standard deviation $\sigma = 1mm$ as radius transition prior $p(r_t|r_{t-1})$. The likelihood model was left unchanged. The application to uncalibrated acquisitions such as MRA would require slightly more in-depth modifications of our Bayesian model. In particular, our likelihood observation model can arguably be improved as it is currently highly dependent on CTA calibration. One lead would be the introduction of a contrast parameter as part of the model. The likelihood formulation could then be split between local contrast estimation and circular geometric fit. This approach would also require the adaptation of the `MFlux` feature, to separate contrast-related information (the norm of the gradient vectors) and geometric information (their direction, supposedly aligned with the model radial directions).

9.3 Publications and Provisional Patent Applications

Journal publication

- Lesage, D., Angelini, A., Bloch, I., Funka-Lea, G., 2009a. A review of 3d vessel lumen segmentation techniques: Models, features and extraction schemes. Med. Image Anal. (in press).

Conference publications

- Lesage, D., Angelini, E., Bloch, I., Funka-Lea, G., 2008. Medial-based bayesian tracking for vascular segmentation: Application to coronary arteries in 3D CT an-

giography. In: Proc. IEEE Int. Symp. Biom. Imaging. pp. 268-271.
– Lesage, D., Angelini, E., Bloch, I., Funka-Lea, G., 2009b. Bayesian maximal paths for coronary artery segmentation from 3D CT angiograms. In: Proc. Med. Image Comput. Comput. Assist. Interv. (in press).
– Lesage, D., Angelini, E., Bloch, I., Funka-Lea, G., 2009c. Design and study of flux-based features for 3d vascular tracking. In: Proc. IEEE Int. Symp. Biom. Imaging. pp. 286- 289.

US provisional patent applications

– **2005P13631US**: Coronary Arteries Segmentation on CT Cardiac Volumes. David Lesage and Matthias Rasch.
– **2009P05215US**: Design and Study of Flux-based Features for 3D Vascular Tracking. David Lesage, Elsa Angelini, Isabelle Bloch, Gareth Funka-Lea.
– **2009P05226US**: Bayesian Maximal Paths for Coronary Artery Segmentation. David Lesage, Elsa Angelini, Isabelle Bloch and Gareth Funka-Lea.

Appendix A

Discretization of the Orientation Domain: Geodesic Spheres

The discretization of a 3D orientation domain can be formulated as the problem of distributing points on a sphere's surface. Unfortunately, there is no known, general analytical solution to the uniform repartition of an arbitrary number of points on a 3D sphere (Wenninger, 1999). There are actually several similar but slightly different ways of defining a "good" repartition of points on a sphere, in terms of *packing*, *covering* or *volume* formulations of the problem (Hardin et al., 2000). One can note that a naive solution, which would consist in defining a set of points with regards to two rotational angles, generates a concentration of points at the poles. Practical solutions to obtain sensible pseudo-uniform sphere distributions include stochastic sampling, simulation techniques based on electro-static repulsion and geometric algorithms such as (Saff and Kuijlaars, 1997), used for vessel segmentation in (Schaap et al., 2007a,b). An alternative is the use of *geodesic spheres*, which offer regular pseudo-uniform layouts (Wenninger, 1999).

Geodesic spheres are classically obtained as iterative subdivisions of Platonic solids. Platonic solids are regular convex polyhedrons with specific symmetry and congruence properties, one of which being that their vertices are uniformly distributed on a sphere. They are optimal solutions to the sphere covering problem, but are limited to very low numbers of vertices. There are only five of such solids (including the cube). The most popular Platonic solid used for the construction of geodesic spheres is the icosahedron, depicted in Fig. A.1. It comprises 20 equilateral faces for 12 vertices. So-called class 1 geodesic spheres are obtained by regularly subdividing faces of the icosahedron, adding new points which are re-projected onto the sphere surface. The number of subdivisions of the face edges is referred to as the *frequency* of the geodesic sphere (Wenninger, 1999) (see also Fig. A.2).

By augmenting the frequency of the subdivision, one obtains geodesic spheres with more surface points (vertices), and thus higher angular resolution. One remark is that contrarily to algorithmic methods such as (Saff and Kuijlaars, 1997), one cannot specify an arbitrary number of surface points, as it is constrained by the subdivision scheme. More precisely, the number of surface points is $10F^2 + 2$, with F the frequency of subdivision. Table A.2 gives the characteristics of class 1 geodesic spheres in function of their subdivision frequency. Another valuable resource on geodesic spheres, including more detailed tables and generation code, is (Hardin et al., 2000).

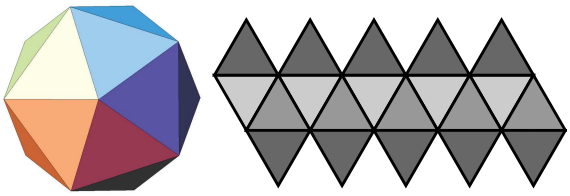


Figure A.1: Icosahedron. Left: 3D rendering. Right: flattened view of the 20 equilateral faces composing the icosahedron. In order to construct geodesic spheres, each face is subdivided, as illustrated in Fig. A.2. In practice, one can focus on a single face, construct the so-called "symmetry triangle", and recover the full 3D geometry by exploiting symmetries.

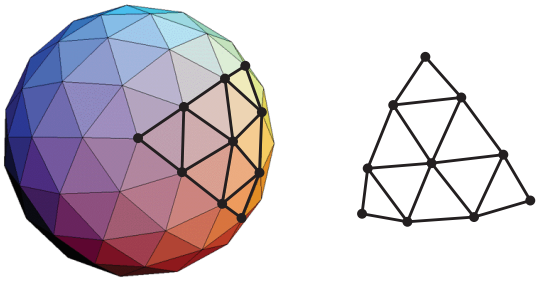


Figure A.2: Example of geodesic sphere (class 1, frequency 3).
Left: 3D rendering. Right: detail of a so-called *symmetry triangle* facet, corresponding to the subdivision of a single face of the original icosahedron.

Table A.1: Characteristics of class 1 geodesic spheres with increasing frequency.
The angular distance corresponds to the maximal angle between a vertex and its closest neighbor. The spatial distance corresponds to the maximal distance between a vertex and its closest neighbor, given for a unit sphere of radius 1. The number of vertices is given by $10F^2 + 2$, with F the frequency.

Frequency F	Number of vertices	Angular resolution (rad.)	Spatial distance $d(F)$ (unit sphere, radius 1)
1	12	1.107	1.051
2	42	0.554	0.547
3	92	0.415	0.412
4	162	0.3	0.299
5	252	0.246	0.245
6	362	0.217	0.217
7	492	0.182	0.182
8	642	0.161	0.16
9	812	0.146	0.146
10	1002	0.13	0.13

Appendix B

Ground Truth Database

B.1 Cardiac CTA Database

The database at our disposal is composed of approximately 150 cardiac CTA datasets, acquired with modern 64-slice CT scanners (Siemens Somatom Sensation 64 and dual source Somatom Definition). The tube voltage used was 120 kV in all cases. Acquisitions were performed in the classical window of maximal arterial opacification, *i.e.* phases ranging from 20% to 80% (the vast majority around 70%). Reconstruction kernels range from B25f to B30f. The average resolution is $0.33 \times 0.33 \times 0.4mm^3$ per voxel. The characteristics of our data are comparable to that of the Rotterdam Evaluation Framework Metz et al. (2008b); Schaap et al. (2009a).

B.2 Manual Ground Truth Delineation

61 datasets were randomly selected for manual delineation:
- 14 Siemens Somatom Definition datasets;
- 47 Siemens Somatom Sensation datasets.

These datasets cover a representative range of image quality, acquisition artifacts and pathologies (soft and calcified plaques, presence of stents). Having a medical expert precisely categorizing our database in terms of image quality and severity of pathologies, as done for instance in the Rotterdam Evaluation Framework, is a future perspective. Manual delineation was carried out by two different persons:
- 19 datasets were initially processed by myself;
- 42 datasets were later delineated by a medical expert, GuangWei Du, M.D, radiologist.

All manual segmentations were cross-validated by both participants.

Particular attention was given to the exhaustiveness of the delineations. We tried, as much as possible, to include all coronary branches we could distinguish. As depicted in Fig. B.1, manual delineation was carried out by specifying successive centerline points inside the vessel lumen, along each coronary branch. We used a limited spacing of successive points (maximum of approximately 5 voxels). The procedure was assisted by semi-automatic tools, extracting notably the cross-sectional contour for each manually specified centerline point. Centerline positions were refined accordingly, as the center of mass of the cross-sectional contour. Manual contour segmentation and correction of the orientation were employed in cases where the results of the automatic contour segmentation procedure were

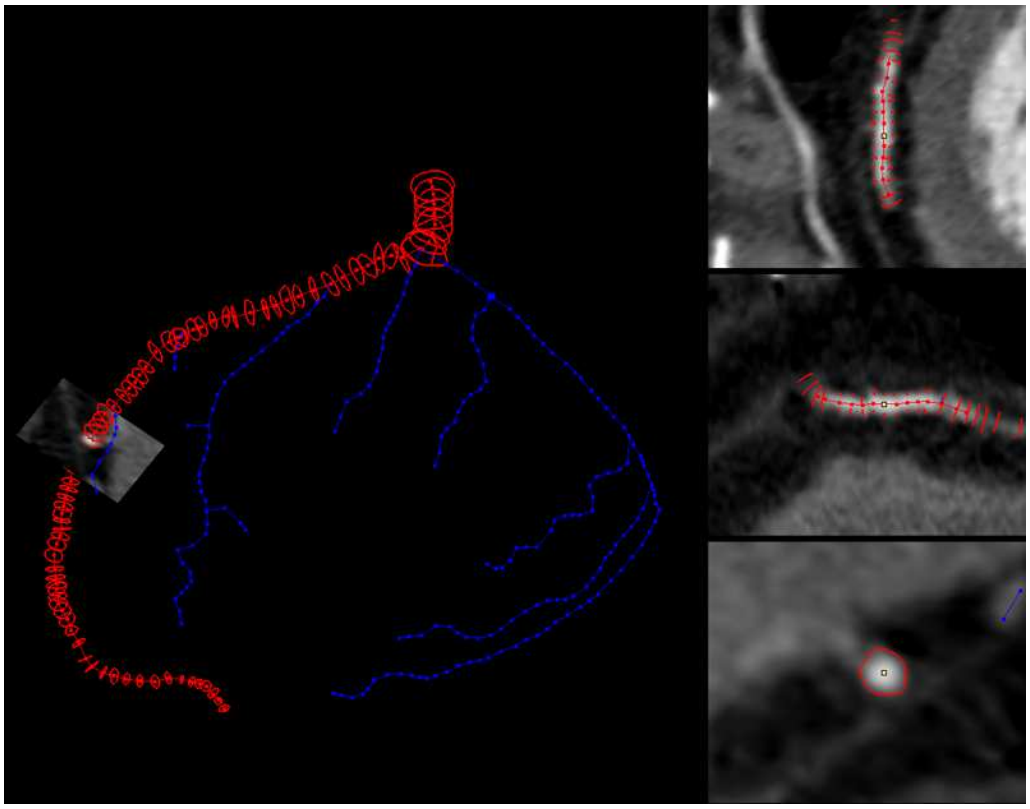


Figure B.1: Manual delineation of coronary trees, obtained by successively specifying centerline locations in the vessel lumen. Corresponding cross-sectional contours are obtained automatically. Left: 3D view of a full coronary tree, with one highlighted branch (in red). Right, from top to bottom: cross-sectional and axial MPR views along the delineated branch.

judged unsatisfactory. Various samples of ground truth delineations from our database can be found throughout this document, *e.g.* in Chapter 7 and 8 (Fig. 7.26, 7.27, 8.21 and 8.22)

For each dataset, both left and right coronary trees were processed. Our ground truth database totals 858 delineated branches, for which centerline and contour (thus average radius) information is available. Manual processing time ranged from 15 to 45 minutes per coronary tree, depending on the complexity of the coronary network at hand.

Compared to the Rotterdam Evaluation Framework, our database comprises more datasets (61 *versus* 32) and more exhaustive delineations (complete trees *versus* 4 main branches). One advantage of the Rotterdam database however is the availability of measures of inter-operator variability, as each of its datasets was delineated by three different experts. Extending our evaluation framework in this direction is a future perspective.

Appendix C

Development and Test Environment

C.1 RadBuilder/XIP Environment

All the algorithms developed during this thesis were implemented in C++, as modules of the RadBuilder / eXtensible Imaging Platform (XIP) (XIP, 2009; Paladini and Azar, 2009). RadBuilder is a modular platform developed internally at Siemens Corporate Research for the rapid prototyping of medical imaging applications. XIP is the Open Source extension of this framework, developed in collaboration with the Mallinckrodt Institute of Radiology at Washington University. The XIP project is part of the Cancer Biomedical Informatics Grid (caBIG) initiative of the US National Institutes of Health (NIH).

The XIP environment offers an extensible set of modules to develop and evaluate software solutions for research and clinical problems. This includes image processing and visualization tools such as volume rendering engines and wrappers for the ITK and VTK libraries. In this framework, researchers can directly focus on the development of their algorithms as isolated modules. These modules can then be integrated within existing networks, including visualization and input manipulation (see Fig. C.1). The modularity of this environment ensures the re-usability of the different tools developed and the rapid deployment of new prototype applications.

In addition to the main algorithms detailed in this manuscript (Chapters 7 and 8), we developed a variety of tools which are currently being re-used in different projects from Siemens Corporate Research. This includes:

- a series of mathematical morphological tools;
- tools for the manipulation of binary masks;
- visualization tools such as an optimized marching cube algorithm;
- a fast, modular implicit active contour implementation derived from (Shi and Karl, 2005);
- a library of modules for the manipulation of vascular segmentation results, detailed in the subsequent section.

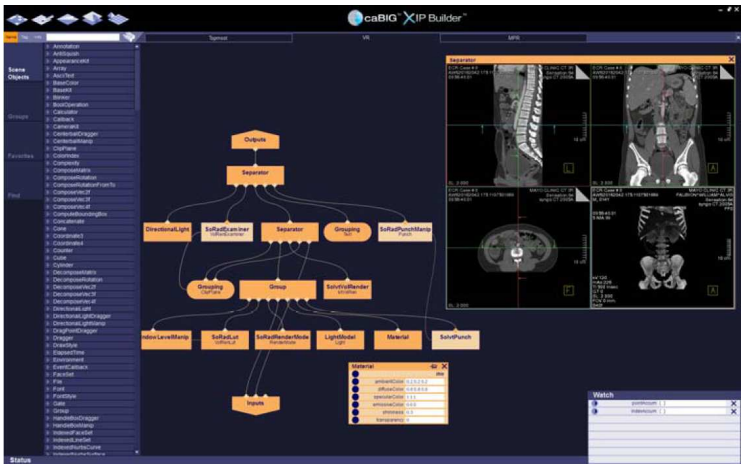


Figure C.1: Screenshot of the XIP Builder environment. The processing pipeline is formed by a modules linked to each others. Modules of interest include readers inputing the data (medical volumes for instance), algorithmic blocks manipulating this data (*e.g.* through image processing operations) and 3D and 2D visualization tools.

```
<vxml>
  <tree label="RCA">
    <branch>
      <cross_section>
        <point3d x="7.12" y="-192.02" z="-214.22"/>
        <radius val="3.14"/>
        <contour>
          ...
        </contour>
      </cross_section>
      ...
    </branch>
    ...
  </tree>
</vxml>
```

Figure C.2: Sample XML result file. A result file can contain several vascular trees, recursively composed of vascular branches. Branches can either be described by binary masks, lists of centerline points, or cross-sections (centerline point + cross-sectional contour).

C.2 Vascular Description Format

We designed an XML format to standardize the description of our vascular segmentation results. An extracted result is described as follows:

- a `vxml` result can contain one or several vascular `tree`s;
- a `tree` is composed recursively of `branch`es;
- a `branch` can be described in one of the following ways:

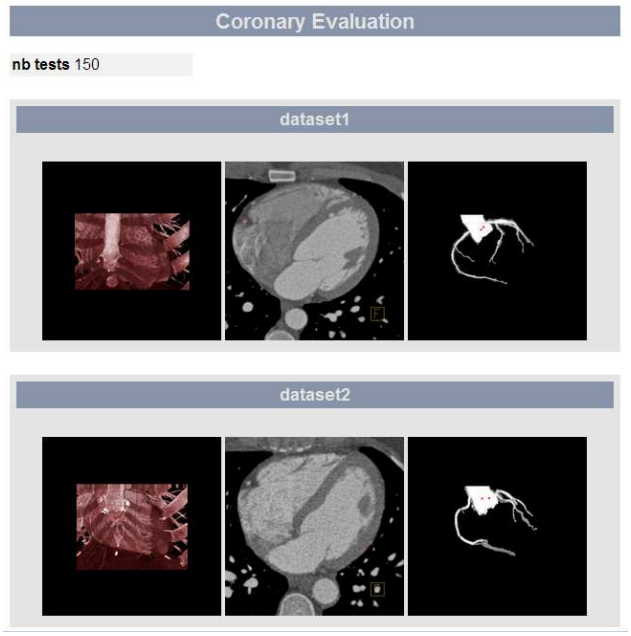


Figure C.3: Sample of automatically generated report. Test scripting can be fully configured to include various screenshots and report quantitative results (computational time, overlap and accuracy measures for instance).

- a binary mask;
- a list of 3D centerline points (`point3d`);
- a 2D cross-section (`cross_section`), composed of a 3D centerline point and a cross-sectional `contour` (list of 3D points).

A sample file is given in Fig. C.2.

This simple but flexible result format has been adopted in different prototypes within Siemens Corporate Research, allowing us to share, compare, extend and correct segmentation results in a direct and natural manner. We implemented a C++ library and corresponding RadBuilder modules to ease the manipulation of such results, including:

- importing and exporting XML results;
- performing graph-based operations such as branch selection and pruning;
- smoothing and interpolating centerline and radius variations;
- translating XML descriptions into OpenGL objects for visualization purposes;
- comparing two result files quantitatively (overlap and accuracy measures, see Sec. 7.7.2.1).

Both our ground truth manual delineations (see Appendix B) and the results extracted by our algorithms share that description format.

C.3 Automatic Testing Environment

Modularity and re-usability were two major concerns of our developments. RadBuilder/XIP provided us with a homogeneous environment in which we conducted all our experiments. On top of this framework, we also built a flexible scripting mechanism for automatic batch testing. Through a system of XML configuration files, our system let

us specify:

- input parameters (*e.g.* dataset input files);
- action scenarios as chains of processing steps (*e.g.* aorta segmentation then ostia detection then coronary extraction);
- output results to be retrieved:
  - segmentation results as XML exports;
  - screenshots of 2D and 3D renderings for qualitative evaluations;
  - quantitative comparisons with respect to the ground truth delineation.

We used this scripting system to perform systematic, qualitative and quantitative evaluations of our work. This way, we only had to specify test scenarios once and were able to apply them automatically on our entire database. For easier monitoring, these scripts can be configured to automatically generate HTML reports, as illustrated in Fig. C.3.

This scripting environment is fully configurable and can be applied to any radBuilder/XIP network. It is currently being re-used in different projects for qualitative and quantitative unit testing.

Appendix D

Minimal Path Extraction Scheme: Rotterdam Evaluation Results

This appendix reports the detailed results obtained by our minimal path algorithm (Chapter 7 of this manuscript) on the Rotterdam Coronary Artery Algorithm Evaluation Framework (formerly known as CAT'08) Schaap et al. (2009a); Metz et al. (2008b).

The evaluation from Metz et al. (2008b) is composed of two sets of data, *training* and *testing* databases. The *training* database comprises 8 datasets and is provided with ground truth segmentation for training and learning purposes. The *testing* database comprises 24 datasets where the ground truth delineation by experts is not provided. Please refer to Schaap et al. (2009a) for details about the evaluation procedure and quantitative measures.

Thereafter, we report detailed *Training* and *Testing* results for the two versions of our minimal path method: interactive (two points) and fully automatic. We additionally include, for comparison, the public results from, to date, 13 other evaluated methods. These public results, along with links to descriptions of the respective methods, are available on the Rotterdam Coronary Artery Algorithm Evaluation Framework website Metz et al. (2008b) and in the associated paper Schaap et al. (2009a).

The Rotterdam framework dissociates three categories of methods (challenges):

- Challenge 1: fully automatic methods;
- Challenge 2: minimally interactive methods (one user-provided point);
- Challenge 3: interactive methods (two or more user-provided points).

Our interactive, two-points method thus falls in Challenge 3 and our fully automatic approach in Challenge 1.

Note on rankings The ranks indicated for our method are *virtual* ranks. As quoted from Metz et al. (2008b):

> *The rank shown here is the rank this submission would get if it would be included in the current ranking. Please see the 'Publishing results' section on the about page for a description on how to include your results in the public results section of this website.*

The results on both the training and testing Rotterdam databases are consistent with our internal validation (see Chapter 7). Compared to other publicly ranked techniques, our approach fares particularly well in terms of overlap robustness (OV, OF and OT measures). Only the technique from (Friman et al., 2008b) outranks our interactive implementation. It should be noted that Friman et al. (2008b) sometimes exploited intermediate seed points

where we restrained ourselves to the provided beginning and end points. In its category, our fully automatic algorithm is outranked solely by the method of Bauer and Bischof (2008a) (average OV measure of 92.7 against 90.2 for our approach) which on the other hand exhibits lower accuracy and higher computational cost.

In terms of centerline accuracy, our algorithms are outranked by the techniques of (Friman et al., 2008b),(Zambal et al., 2008),(Szymczak, 2008) and (Kitslaar et al., 2008). One can note that the approaches of (Friman et al., 2008b) and (Zambal et al., 2008) employ local refinement steps for increased accuracy. Similarly, the accuracy of our centerline extraction could be improved by subsequent post-processing steps such as an accurate cross-sectional contour segmentation (see Sec. 9.2). Without such additional processing, the accuracy obtained is of the order of the data resolution (approximately 0.3mm of average distance AI). This result is in accordance with our primary objective, which was to favor delineation robustness over accuracy. More detailed discussions of our evaluation results can be found in Chapter 7.

The results reported in this appendix will be shortly made available online on the website of (Metz et al., 2008b), including the final ranking of our algorithms within other methods. We invite the reader to refer to (Metz et al., 2008b) for up-to-date results and rankings. Additionally, the Rotterdam framework website proposes convenient interactive tools for comparing the results of the different algorithms.

Table D.1: Interactive workflow: results summary on the training database (8 datasets).

Measure	% / mm			score			rank		
	min.	max.	avg.	min.	max.	avg.	min.	max.	avg.
OV	92.1%	100.0%	98.8%	46.4	100.0	88.5	1	4	1.75
OF	18.9%	100.0%	88.1%	9.6	100.0	81.6	1	8	2.12
OT	92.1%	100.0%	99.0%	46.3	100.0	86.7	1	5	1.78
AI	0.22 mm	0.47 mm	0.31 mm	27.4	47.6	37.5	3	7	4.75
Total							**1**	**8**	**3.32**

Table D.2: Interactive workflow: results summary on the testing database (24 datasets).

Measure	% / mm			score			rank		
	min.	max.	avg.	min.	max.	avg.	min.	max.	avg.
OV	74.4%	100.0%	97.6%	38.9	100.0	81.3	1	8	2.32
OF	5.3%	100.0%	78.6%	2.6	100.0	70.4	1	14	3.09
OT	74.6%	100.0%	97.7%	38.8	100.0	81.3	1	8	2.40
AI	0.21 mm	0.54 mm	0.29 mm	26.7	57.0	36.9	1	11	5.19
Total							**1**	**14**	**3.90**

Table D.3: Interactive workflow, training database: average overlap per dataset

Dataset	OV			OF			OT			Avg.
nr.	%	score	rank	%	score	rank	%	score	rank	rank
0	97.6	77.8	2.50	79.7	76.6	3.00	97.6	77.5	2.50	2.67
1	99.8	87.4	1.25	98.2	86.6	1.50	99.8	87.4	1.25	1.32
2	100.0	100.0	1.00	100.0	100.0	1.00	100.0	100.0	1.00	1.00
3	97.8	92.9	2.00	89.2	82.8	2.00	97.8	92.2	2.00	2.00
4	99.2	90.4	2.00	80.5	76.3	2.00	99.4	87.5	1.50	1.85
5	99.3	85.5	2.25	96.0	74.2	2.50	99.3	74.8	2.50	2.42
6	100.0	100.0	1.00	100.0	100.0	1.00	100.0	100.0	1.00	1.00
7	97.0	73.7	2.00	61.0	56.4	4.00	97.9	74.0	2.50	2.83
Avg.	**98.8**	**88.5**	**1.75**	**88.1**	**81.6**	**2.12**	**99.0**	**86.7**	**1.78**	**1.89**

Table D.4: Interactive workflow, training database: average accuracy per dataset

Dataset	AI			Avg.
nr.	mm	score	rank	rank
0	0.40	38.1	4.50	4.50
1	0.31	37.6	5.25	5.25
2	0.27	35.7	4.50	4.50
3	0.31	40.5	5.00	5.00
4	0.24	38.9	4.75	4.75
5	0.37	38.9	5.50	5.50
6	0.26	36.2	4.00	4.00
7	0.32	34.3	4.50	4.50
Avg.	**0.31**	**37.5**	**4.75**	**4.75**

Table D.5: Interactive workflow, testing database: average overlap per dataset

Dataset	OV			OF			OT			Avg.
nr.	%	score	rank	%	score	rank	%	score	rank	rank
8	91.9	74.7	2.50	79.6	73.2	2.25	93.3	74.9	3.00	2.60
9	99.9	94.5	1.75	99.7	97.6	2.00	99.9	87.5	2.00	1.93
10	98.5	81.0	2.50	96.1	74.1	3.00	98.9	87.5	1.25	2.25
11	95.9	61.6	2.75	53.8	53.1	4.25	95.9	61.5	2.75	3.25
12	97.8	78.5	2.25	50.9	39.9	3.25	97.9	74.3	2.50	2.67
13	98.8	69.8	3.75	73.1	49.8	5.75	98.9	80.0	3.75	4.40
14	99.0	87.1	2.00	83.9	82.6	2.50	99.0	87.1	2.25	2.25
15	100.0	100.0	1.00	100.0	100.0	1.00	100.0	100.0	1.00	1.00
16	99.9	94.4	1.25	100.0	100.0	1.00	100.0	100.0	1.00	1.07
17	93.5	60.4	4.50	45.5	36.1	7.75	93.5	60.0	5.00	5.73
18	94.7	73.4	3.25	79.7	65.3	2.75	94.7	73.4	3.25	3.10
19	100.0	100.0	1.00	100.0	100.0	1.00	100.0	100.0	1.00	1.00
20	97.8	67.5	3.75	53.1	41.8	5.50	97.8	61.8	4.00	4.40
21	100.0	100.0	1.00	100.0	100.0	1.00	100.0	100.0	1.00	1.00
22	100.0	100.0	1.00	100.0	100.0	1.00	100.0	100.0	1.00	1.00
23	99.3	84.3	3.25	97.4	74.3	3.00	99.3	80.1	3.25	3.15
24	96.5	85.9	3.00	75.7	71.8	4.25	96.6	85.6	3.50	3.58
25	98.9	62.2	1.25	53.5	39.7	3.50	99.1	62.1	2.75	2.47
26	86.1	57.4	2.50	34.7	32.8	6.25	86.0	57.2	2.50	3.75
27	96.3	74.4	3.00	64.0	53.9	2.00	96.2	66.0	3.25	2.75
28	97.9	79.0	3.00	74.6	64.0	5.00	97.8	74.3	3.25	3.75
29	99.1	74.7	2.50	77.4	64.3	3.25	99.5	87.3	1.00	2.25
30	99.6	96.2	1.25	96.0	87.4	1.25	99.8	95.5	1.25	1.25
31	99.9	95.3	1.75	98.0	86.8	1.75	99.9	94.5	2.00	1.82
Avg.	**97.6**	**81.3**	**2.32**	**78.6**	**70.4**	**3.09**	**97.7**	**81.3**	**2.40**	**2.60**

Table D.6: Interactive workflow, testing database: accuracy per dataset

Dataset	AI			Avg.
nr.	mm	score	rank	rank
8	0.34	42.6	4.75	4.75
9	0.24	35.6	5.50	5.50
10	0.31	33.6	5.50	5.50
11	0.31	41.6	3.75	3.75
12	0.28	36.5	4.25	4.25
13	0.27	38.3	5.75	5.75
14	0.34	38.2	6.75	6.75
15	0.28	37.2	5.75	5.75
16	0.27	34.9	4.25	4.25
17	0.44	36.6	8.00	8.00
18	0.26	36.9	5.50	5.50
19	0.31	40.6	5.25	5.25
20	0.40	35.1	6.75	6.75
21	0.23	35.1	5.50	5.50
22	0.26	37.9	4.25	4.25
23	0.28	37.4	5.25	5.25
24	0.22	34.9	5.00	5.00
25	0.27	38.4	2.50	2.50
26	0.38	47.6	5.00	5.00
27	0.34	37.9	4.25	4.25
28	0.24	31.5	5.25	5.25
29	0.27	33.0	5.25	5.25
30	0.24	35.1	4.75	4.75
31	0.24	28.7	5.75	5.75
Avg.	**0.29**	**36.9**	**5.19**	**5.19**

Table D.7: Automatic workflow: results summary on the training database (8 datasets).

Measure	% / mm			score			rank		
	min.	max.	avg.	min.	max.	avg.	min.	max.	avg.
OV	57.3%	99.0%	90.2%	32.0	91.5	50.5	3	10	6.28
OF	20.5%	100.0%	81.0%	13.1	100.0	54.2	1	8	4.81
OT	57.3%	100.0%	93.2%	33.0	100.0	61.7	1	9	5.06
AI	0.22 mm	0.46 mm	0.33 mm	27.5	49.8	36.9	2	7	4.91
Total							**1**	**10**	**5.15**

Table D.8: Automatic workflow: results summary on the testing database (24 datasets).

Measure	% / mm			score			rank		
	min.	max.	avg.	min.	max.	avg.	min.	max.	avg.
OV	58.9%	99.7%	91.3%	29.5	96.7	50.9	1	14	6.94
OF	5.3%	100.0%	75.3%	2.6	100.0	55.7	1	14	4.69
OT	58.9%	100.0%	93.9%	29.5	100.0	67.7	1	14	4.67
AI	0.21 mm	0.56 mm	0.31 mm	24.1	61.5	36.3	2	10	5.52
Total							**1**	**14**	**5.48**

Table D.9: Automatic workflow, training database: average overlap per dataset

Dataset	OV			OF			OT			Avg.
nr.	%	score	rank	%	score	rank	%	score	rank	rank
0	94.4	48.7	5.50	57.8	38.4	7.00	95.1	49.0	5.50	5.97
1	94.5	47.4	6.75	91.9	46.2	6.00	94.8	47.5	6.75	6.50
2	93.7	48.3	8.00	95.4	65.4	4.50	98.9	66.9	5.25	5.92
3	75.5	41.2	8.00	60.2	50.4	4.75	79.1	59.4	5.75	6.17
4	94.7	64.8	5.00	90.9	62.6	3.50	98.2	74.3	3.50	4.00
5	93.9	52.5	5.00	94.5	73.4	3.00	97.1	73.6	3.50	3.83
6	93.1	56.2	6.00	87.7	50.7	5.00	94.3	75.0	4.25	5.08
7	81.9	45.2	6.00	69.9	46.5	4.75	87.7	47.7	6.00	5.58
Avg.	**90.2**	**50.5**	**6.28**	**81.0**	**54.2**	**4.81**	**93.2**	**61.7**	**5.06**	**5.38**

Table D.10: Automatic workflow, training database: average accuracy per dataset

Dataset	AI			Avg.
nr.	mm	score	rank	rank
0	0.39	40.2	4.50	4.50
1	0.34	35.9	5.75	5.75
2	0.27	36.1	4.50	4.50
3	0.34	41.7	4.75	4.75
4	0.26	35.2	5.00	5.00
5	0.36	39.3	4.75	4.75
6	0.32	33.0	4.50	4.50
7	0.37	33.5	5.50	5.50
Avg.	**0.33**	**36.9**	**4.91**	**4.91**

Table D.11: Automatic workflow, testing database: average overlap per dataset

Dataset	OV			OF			OT			Avg.
nr.	%	score	rank	%	score	rank	%	score	rank	rank
8	84.5	59.5	6.00	56.7	49.5	4.75	87.3	61.0	5.00	5.25
9	94.4	55.4	6.00	88.0	61.8	5.75	96.7	71.1	5.00	5.60
10	87.4	49.6	7.00	81.8	54.4	2.75	88.6	57.9	4.75	4.85
11	91.8	46.9	6.00	53.6	53.1	4.00	92.4	59.7	4.50	4.83
12	94.5	51.3	6.00	45.7	36.9	5.50	97.0	63.8	3.50	5.03
13	96.9	49.1	6.00	96.2	74.1	2.75	98.9	87.2	2.00	3.60
14	93.8	47.5	7.00	83.6	69.9	3.50	99.0	87.1	2.00	4.17
15	89.7	54.6	10.00	85.7	68.8	5.00	94.6	76.8	6.00	7.00
16	93.2	51.9	7.25	93.3	69.3	5.50	97.5	79.3	5.00	5.95
17	90.9	46.8	6.00	65.2	33.8	7.00	94.6	49.5	5.25	6.08
18	90.9	46.2	8.25	75.8	59.5	5.75	92.1	59.0	5.50	6.50
19	97.2	61.0	6.75	100.0	100.0	1.00	100.0	100.0	1.00	2.92
20	91.4	50.5	7.75	50.8	28.1	6.25	92.9	46.8	7.25	7.08
21	96.1	54.6	7.50	95.3	75.0	6.75	98.1	74.6	5.50	6.60
22	97.2	49.0	6.75	97.6	61.3	4.50	99.3	74.6	4.00	5.10
23	94.4	47.7	9.00	94.5	60.3	5.00	95.7	60.5	7.25	7.10
24	89.7	46.2	7.25	69.5	60.3	4.75	93.1	67.4	4.75	5.58
25	97.7	49.7	3.00	69.9	55.5	2.50	99.2	74.6	1.50	2.33
26	84.0	44.0	3.75	40.7	37.5	1.50	85.4	56.8	3.25	2.83
27	69.6	36.8	9.00	47.3	29.1	5.00	71.9	37.2	9.00	7.65
28	95.1	58.3	6.25	73.7	48.5	7.25	97.0	73.8	5.00	6.17
29	90.5	45.5	8.00	67.2	34.3	6.50	93.9	71.9	4.50	6.33
30	88.0	49.5	8.50	83.4	43.8	6.50	91.3	48.5	8.00	7.65
31	93.3	68.9	7.50	92.0	71.2	2.75	96.2	85.7	2.50	4.25
Avg.	**91.3**	**50.9**	**6.94**	**75.3**	**55.7**	**4.69**	**93.9**	**67.7**	**4.67**	**5.43**

Table D.12: Automatic workflow, testing database: average accuracy per dataset

Dataset	AI			Avg.
nr.	mm	score	rank	rank
8	0.35	43.4	5.25	5.25
9	0.25	35.4	5.25	5.25
10	0.33	33.6	5.50	5.50
11	0.33	40.2	4.50	4.50
12	0.31	34.6	5.25	5.25
13	0.26	39.3	5.25	5.25
14	0.33	38.6	6.50	6.50
15	0.28	37.5	5.50	5.50
16	0.28	34.2	5.25	5.25
17	0.41	36.8	8.00	8.00
18	0.28	34.7	6.50	6.50
19	0.30	42.4	4.75	4.75
20	0.43	34.6	7.50	7.50
21	0.26	33.8	5.50	5.50
22	0.26	38.2	3.75	3.75
23	0.28	37.7	5.50	5.50
24	0.23	33.6	5.00	5.00
25	0.30	37.3	3.00	3.00
26	0.42	46.7	5.50	5.50
27	0.40	34.9	6.00	6.00
28	0.25	32.0	5.50	5.50
29	0.30	30.0	6.75	6.75
30	0.25	33.7	5.00	5.00
31	0.24	29.0	6.00	6.00
Avg.	**0.31**	**36.3**	**5.52**	**5.52**

Table D.13: Training (8 datasets). The overall ranking of the 13 evaluated methods. Challenge 1: fully automatic methods. Challenge 2: minimally interactive methods (one user-provided point). Challenge 3: interactive methods (two or more user-provided points).

Method	**Challenge**			**Avg. Ov. rank**	**Avg. Acc. rank**	**Avg. rank**
	1	**2**	**3**			
Friman et al. - MHT			×	1.62	1.26	1.44
Szymczak - Tracer			×	3.88	2.68	3.28
Krissian et al. - KnowledgeBasedMinPath		×		1.94	5.61	3.78
Zambal et al. - DepthFirstModelFit	×			4.60	3.30	3.95
Metz et al. - TwoPointMinCost			×	3.79	6.35	5.07
Kitslaar et al. - CocomoBeach	×			6.52	4.20	5.36
Bauer and Bischof - GVFTube'n'Linkage	×			5.53	5.67	5.60
Zhang et al. - 3DInteractiveTrack			×	5.93	7.28	6.61
Castro et al. - CoronaryTreeMorphoRec		×		7.56	8.65	8.11
Wang and Smedby - VirtualContrast	×			N/A	N/A	N/A
Dikici et al. - AxialSymmetry		×		N/A	N/A	N/A
Tek et al. - AutoCoronaryTree	×			N/A	N/A	N/A
Hernández Hoyos et al. - ElasticModel			×	N/A	N/A	N/A

Table D.14: Testing (24 datasets). The overall ranking of the 13 evaluated methods. Challenge 1: fully automatic methods. Challenge 2: minimally interactive methods (one user-provided point). Challenge 3: interactive methods (two or more user-provided points).

Method	**Challenge**			**Avg. Ov. rank**	**Avg. Acc. rank**	**Avg. rank**
	1	**2**	**3**			
Friman et al. - MHT			×	2.07	1.42	1.75
Szymczak - Tracer			×	4.21	2.35	3.28
Zambal et al. - DepthFirstModelFit	×			6.17	3.82	5.00
Krissian et al. - KnowledgeBasedMinPath		×		4.31	7.94	6.13
Bauer and Bischof - GVFTube'n'Linkage	×			5.39	7.29	6.34
Tek et al. - AutoCoronaryTree	×			7.69	5.74	6.72
Metz et al. - TwoPointMinCost			×	5.30	8.17	6.74
Kitslaar et al. - CocomoBeach	×			8.56	5.75	7.16
Dikici et al. - AxialSymmetry		×		6.95	9.11	8.03
Wang and Smedby - VirtualContrast	×			8.71	8.36	8.54
Zhang et al. - 3DInteractiveTrack			×	7.52	10.33	8.93
Hernández Hoyos et al. - ElasticModel			×	9.05	8.91	8.98
Castro et al. - CoronaryTreeMorphoRec		×		10.42	11.77	11.10

Table D.15: Training (8 datasets). Overlap measures for the 13 evaluated methods. The average overlap, score and rank is shown for each of the three overlap measures. Challenge 1: fully automatic methods. Challenge 2: minimally interactive methods (one user-provided point). Challenge 3: interactive methods (two or more user-provided points).

Method	**Challenge**			**OV**			**OF**			**OT**		
	1	**2**	**3**	**%**	**score**	**rank**	**%**	**score**	**rank**	**%**	**score**	**rank**
Friman et al. - MHT			×	99.3	91.6	1.66	94.6	83.2	1.78	99.5	90.0	1.41
Szymczak - Tracer			×	92.2	68.5	3.50	68.0	52.9	4.56	92.8	68.3	3.59
Krissian et al. - KnowledgeBasedMinPath		×		98.1	88.1	1.88	92.7	85.3	2.00	98.1	89.4	1.94
Zambal et al. - DepthFirstModelFit	×			89.2	52.2	5.75	78.3	60.3	4.03	93.0	67.3	4.03
Metz et al. - TwoPointMinCost			×	93.3	71.2	3.63	67.8	57.8	4.16	93.5	72.5	3.59
Kitslaar et al. - CocomoBeach	×			80.2	44.3	7.19	66.4	41.7	5.50	81.6	48.1	6.88
Bauer and Bischof - GVFTube'n'Linkage	×			89.6	53.0	5.94	66.6	45.3	5.38	91.5	58.4	5.28
Zhang et al. - 3DInteractiveTrack			×	91.6	54.2	5.53	57.9	36.1	6.38	92.0	51.8	5.88
Castro et al. - CoronaryTreeMorphoRec		×		72.6	38.9	7.75	45.6	27.3	7.25	73.8	41.0	7.69
Wang and Smedby - VirtualContrast	×			0.0	0.0	10.00	0.0	0.0	10.00	0.0	0.0	10.00
Dikici et al. - AxialSymmetry		×		0.0	0.0	10.00	0.0	0.0	10.00	0.0	0.0	10.00
Tek et al. - AutoCoronaryTree	×			0.0	0.0	10.00	0.0	0.0	10.00	0.0	0.0	10.00
Hernández Hoyos et al. - ElasticModel			×	0.0	0.0	10.00	0.0	0.0	10.00	0.0	0.0	10.00

Table D.16: Testing (24 datasets). Overlap measures for the 13 evaluated methods. The average overlap, score and rank is shown for each of the three overlap measures. Challenge 1: fully automatic methods. Challenge 2: minimally interactive methods (one user-provided point). Challenge 3: interactive methods (two or more user-provided points).

Method	**Challenge**			**OV**			**OF**			**OT**		
	1	**2**	**3**	**%**	**score**	**rank**	**%**	**score**	**rank**	**%**	**score**	**rank**
Friman et al. - MHT			×	98.5	84.0	1.74	83.1	72.8	2.64	98.7	84.5	1.83
Szymczak - Tracer			×	95.1	71.0	3.60	63.5	52.0	5.22	95.5	70.2	3.81
Zambal et al. - DepthFirstModelFit	×			84.7	48.6	7.29	65.3	49.2	5.32	87.0	60.1	5.90
Krissian et al. - KnowledgeBasedMinPath		×		88.0	67.4	4.46	74.2	61.1	4.27	88.5	70.0	4.21
Bauer and Bischof - GVFTube'n'Linkage	×			92.7	52.3	6.20	71.9	51.4	5.32	95.3	67.0	4.66
Tek et al. - AutoCoronaryTree	×			84.7	46.5	8.13	59.5	36.1	7.26	86.2	50.3	7.69
Metz et al. - TwoPointMinCost			×	91.9	64.5	4.70	56.4	45.6	6.22	92.5	64.5	4.97
Kitslaar et al. - CocomoBeach	×			78.8	42.5	9.34	64.4	40.0	7.39	81.2	46.9	8.96
Dikici et al. - AxialSymmetry		×		90.8	56.8	6.17	48.9	35.6	7.96	91.7	55.9	6.71
Wang and Smedby - VirtualContrast	×			75.6	39.2	9.74	56.1	34.5	7.74	78.7	45.6	8.64
Zhang et al. - 3DInteractiveTrack			×	89.6	51.1	7.04	49.9	30.5	8.36	90.6	52.4	7.15
Hernández Hoyos et al. - ElasticModel			×	77.0	40.5	9.60	52.1	31.5	8.46	79.0	45.3	9.09
Castro et al. - CoronaryTreeMorphoRec		×		67.0	34.5	11.00	36.3	20.5	9.53	69.1	36.7	10.74

Table D.17: Training (8 datasets). Accuracy measures for the 13 evaluated methods. The average distance, score and rank of each is shown for the accuracy when inside measure (AI). Challenge 1: fully automatic methods. Challenge 2: minimally interactive methods (one user-provided point). Challenge 3: interactive methods (two or more user-provided points).

Method	**Challenge**			**AD**			**AI**			**AT**		
	1	**2**	**3**	**mm**	**score**	**rank**	**mm**	**score**	**rank**	**mm**	**score**	**rank**
Friman et al. - MHT			×	0.25	47.6	1.16	0.24	47.9	1.41	0.25	47.8	1.22
Szymczak - Tracer			×	0.60	40.8	2.75	0.28	43.8	2.56	0.59	41.1	2.72
Krissian et al. - KnowledgeBasedMinPath		×		0.44	30.0	5.34	0.41	30.4	6.03	0.44	29.9	5.47
Zambal et al. - DepthFirstModelFit	×			1.72	36.8	3.44	0.31	40.5	3.25	1.14	38.1	3.22
Metz et al. - TwoPointMinCost			×	0.71	27.0	6.19	0.48	28.5	6.59	0.71	26.9	6.28
Kitslaar et al. - CocomoBeach	×			3.85	33.0	4.56	0.32	40.6	3.41	3.60	33.4	4.63
Bauer and Bischof - GVFTube'n'Linkage	×			1.62	28.4	5.72	0.39	31.3	5.63	1.48	28.8	5.66
Zhang et al. - 3DInteractiveTrack			×	0.63	24.3	7.16	0.51	25.8	7.53	0.63	24.3	7.16
Castro et al. - CoronaryTreeMorphoRec		×		5.60	16.0	8.69	0.67	21.2	8.59	5.41	16.2	8.66
Wang and Smedby - VirtualContrast	×			0.00	0.0	10.00	0.00	0.0	10.00	0.00	0.0	10.00
Dikici et al. - AxialSymmetry		×		0.00	0.0	10.00	0.00	0.0	10.00	0.00	0.0	10.00
Tek et al. - AutoCoronaryTree	×			0.00	0.0	10.00	0.00	0.0	10.00	0.00	0.0	10.00
Hernández Hoyos et al. - ElasticModel			×	0.00	0.0	10.00	0.00	0.0	10.00	0.00	0.0	10.00

Table D.18: Testing (24 datasets). Accuracy measures for the 13 evaluated methods. The average distance, score and rank of each is shown for the accuracy when inside measure (AI). Challenge 1: fully automatic methods. Challenge 2: minimally interactive methods (one user-provided point). Challenge 3: interactive methods (two or more user-provided points).

Method	**Challenge**			**AD**			**AI**			**AT**		
	1	**2**	**3**	**mm**	**score**	**rank**	**mm**	**score**	**rank**	**mm**	**score**	**rank**
Friman et al. - MHT			×	0.25	47.5	1.34	0.23	47.9	1.58	0.25	47.8	1.34
Szymczak - Tracer			×	0.36	42.7	2.23	0.26	44.4	2.52	0.35	43.0	2.31
Zambal et al. - DepthFirstModelFit	×			3.54	35.7	4.09	0.28	41.9	3.33	3.18	36.5	4.03
Krissian et al. - KnowledgeBasedMinPath		×		3.52	25.8	7.68	0.39	29.2	8.36	3.74	25.9	7.79
Bauer and Bischof - GVFTube'n'Linkage	×			0.74	28.3	7.01	0.37	29.8	8.02	0.54	29.0	6.85
Tek et al. - AutoCoronaryTree	×			3.21	30.7	6.04	0.34	35.3	5.18	2.92	31.3	6.01
Metz et al. - TwoPointMinCost			×	1.09	26.2	7.77	0.46	28.0	8.80	1.08	26.2	7.95
Kitslaar et al. - CocomoBeach	×			5.46	30.8	6.16	0.29	37.7	5.04	4.70	31.6	6.04
Dikici et al. - AxialSymmetry		×		0.85	24.5	8.75	0.46	26.4	9.60	0.81	24.4	8.98
Wang and Smedby - VirtualContrast	×			6.52	23.5	8.77	0.39	30.6	7.74	5.87	24.3	8.57
Zhang et al. - 3DInteractiveTrack			×	0.63	22.5	10.01	0.51	24.2	10.91	0.63	22.6	10.06
Hernández Hoyos et al. - ElasticModel			×	6.07	22.7	9.23	0.40	29.3	8.29	5.53	23.2	9.21
Castro et al. - CoronaryTreeMorphoRec		×		8.46	15.1	11.90	0.59	20.7	11.59	7.94	15.6	11.82

Appendix E

Medial-Based Bayesian Tracking For Vascular Segmentation: Application To Coronary Arteries In 3D CT Angiography

This appendix is the reproduction of (Lesage et al., 2008), published in IEEE International Symposium on Biomedical Imaging (ISBI 2008). This publication corresponds to an early work on the use of particle filtering for the tracking of vascular structures. Following are the key differences with the developments detailed in Chapter 8:

- the approach in (Lesage et al., 2008) was limited to the extraction of a single branch; the algorithm from Chapter 8 can detect and track multiple branches;
- it relied on a slightly different geometric model than the one detailed in this dissertation (Chapter 4);
- in terms of image features, it exploited morphological operators (multiscale erosions) and a multiscale, spherical gradient flux measure; in contrast, the method from Chapter 8 relies on the *oriented* `MFlux` feature for more robust, less noise-sensitive results;
- the radius and direction transition priors in this early work were learned independently of the current radius r_t; in contrast, our extended Bayesian model (Chapter 6) distinguishes subtle differences in radius and direction variations with respect to the current radius (see Fig. 6.6 and 6.8);
- the work in (Lesage et al., 2008) focused on a new sampling scheme, based on the discretization of the posterior distribution of successor candidates; the method in Chapter 8 does not require a discretization of the search space, which has accuracy benefits, and relies on auxiliary simulation rather than the full evaluation of a set of discretized candidates.

E.1 Introduction

In biomedical applications, vascular structures are often of critical importance for diagnosis, treatment and surgery planning. Vessels are thin, elongated and complex structures embedded in increasingly large images. Manual delineation, although still heavily used in clinical routines, has become a considerable burden and automatic or semi-automatic

segmentation remains challenging.

Vascular segmentation has received considerable attention in the literature Kirbas and Quek (2004). A popular approach is to consider the segmentation as an iterative, *tracking* process. Classical region-growing techniques can be seen as primitive representatives of this class of methods. Front propagation techniques allow for refined analysis by imposing a structurally coherent exploration process. The robustness of local deterministic tracking is generally limited by the necessity of using low-level, causal criteria. In some settings, the tracking problem can be formulated as the extraction of globally minimal paths Deschamps and Cohen (2000); Li and Yezzi (2006). Another approach, which is becoming increasingly popular, is the use of stochastic Bayesian tracking algorithms such as particle filters Florin et al. (2005); Schaap et al. (2007a). Such algorithms have demonstrated particular robustness while allowing for high-level modeling.

In this paper, we propose a new Bayesian, stochastic tracking algorithm, inspired by recent developments in particle filtering design. It relies on a constrained, medial-based geometric model and on an original sampling scheme for the selection of tracking hypotheses. It was applied to coronary artery segmentation from 3D cardiac computed tomography (CT) images, a particularly challenging task in terms of anatomical variability and complexity. Prior knowledge was learned from a manually segmented database. Qualitative and quantitative validation is presented on clinical data.

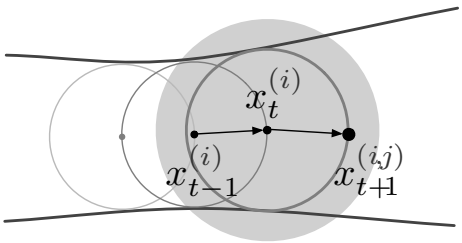


Figure E.1: Geometric model and successor candidates. The gray area indicates the potential successors of $x_t^{(i)}$, e.g. $x_{t+1}^{(i,j)}$.

E.2 Geometric Model

For the representation of vascular structures, we advocate the use of parametric geometric models for their conciseness and usability for further analysis. Our model follows directly from Blum's seminal work on medial representations Blum (1967). A binary object is represented by a collection of maximally included hyper-spheres $\{S_k\}$ of centers $\{x_k\}$ and radiuses $\{r_k\}$. As illustrated in Figure E.1, sphere centers form the *medial axis* of the object and associated radiuses allow for the reconstruction of the object surface Blum (1967).

Our model considers a subset of Blum's representation, defined by a geometric constraint. In our tracking framework, our geometric model is iteratively updated by adding new spheres centers which lie on the surface of the preceding sphere. As depicted in Figure E.1, a new center x_t is at distance r_{t-1} of x_{t-1}. A non-branching structure is then represented by a *discrete* chain of spheres $\{S_0, ..., S_T\}$.

Thanks to this connectivity constraint, our model is entirely described with the col-

lection of centers $\{x_k\}$. The radius r_t and axis direction d_t appear as linked variables:

$$\vec{d_t} = x_{t+1} - x_t \quad \text{and} \quad r_t = |x_{t+1} - x_t|$$

In the following, we reduce all the notations to the centers $\{x_t\}$, keeping in mind that a center at time t sets radius and direction variables at time $t-1$ through our geometric connectivity constraint.

Numerous tracking techniques for tubular structures rely on similar medial-based models. In the vascular segmentation field, one can cite particular efforts for centerline tracking based on cross-section detection Florin et al. (2005); Schaap et al. (2007a); Wink et al. (2000a); Aylward and Bullitt (2002); Fridman et al. (2003). Such techniques incorporate the direction and scale in their state vectors, resulting in large search spaces, especially in 3D. In contrast, our search space is limited to the geometric location of the successive centers, but we are intrinsically limited to tubular structures with *circular* cross-sections. This geometric model directly applies to elongated structures in 2D and 3D images. It can be extended to branching structures by using a tree or graph of spheres instead of a simple chain.

E.3 Tracking Scheme

The proposed tracking scheme is an iterative, population-based, stochastic Bayesian filter, largely inspired by recent developments in particle filtering for segmentation tasks. It aims at recovering a single vascular branch, from its root to its distal end, as a chain of spheres centers $X_T = \{x_0, ..., x_T\}$ following the geometric model described previously and estimated given the observations, *i.e.*, the image, considered as stationary: $Y_t = Y, \forall t$.

Similarly to classical particle filters, we use a first order hidden Markov model and a Monte-Carlo approximation of the posterior distribution $p(X_t|Y)$ by a weighted population of N_t discrete samples $\{x_t^{(i)}, w_t^{(i)}\}_{i=1}^{N_t}$ Doucet et al. (2001). Instead of using classical *update* and *prediction* steps, we introduce a new sampling scheme to recursively evolve this population.

At time-step t, we consider all the potential successors $\{x_{t+1}^{(i,j)}\}$ of $x_t^{(i)}$ as the discrete locations at distances $r_t^{(i,j)} \in [0, r_{max}]$ of $x_t^{(i)}$ (gray area in Figure E.1). Importantly, a candidate $x_{t+1}^{(i,j)}$ *retrospectively* fixes radius $r_t^{(i,j)}$ and direction $\vec{d_t}^{(i,j)}$ according to our model. Associated weights $w_{t+1}^{(i,j)}$ are evaluated as follows:

$$w_{t+1}^{(i,j)} \propto w_t^{(i)} p(Y|x_{t+1}^{(i,j)}) p(x_{t+1}^{(i,j)}|x_t^{(i)})$$

where $p(Y|x_{t+1}^{(i,j)})$ and $p(x_{t+1}^{(i,j)}|x_t^{(i)})$ are the likelihood and dynamic prior of the candidate $x_{t+1}^{(i,j)}$, respectively. The *extended* population $\{x_{t+1}^{(i,j)}, w_{t+1}^{(i,j)}\}$ forms a distribution from which the new population of N_{t+1} samples $\{x_{t+1}^{(i)}, w_{t+1}^{(i)}\}_{i=1}^{N_{t+1}}$ is obtained by *importance* sampling Doucet et al. (2001). When a candidate $(x_{t+1}^{(i,j)}, w_{t+1}^{(i,j)})$ is selected, it enters the new population as $(x_{t+1}^{(i)} = x_{t+1}^{(i,j)}, w_{t+1}^{(i)} = x_{t+1}^{(i,j)})$. Weights are then normalized before recursively proceeding to the next step.

This scheme can be seen as a systematic resampling step with *delayed* application of the dynamic prior. We emphasize that, in contrast with classical particle filters, our algorithm actually takes into account a wider range of hypotheses by evaluating *all* the potential successors $\{x_{t+1}^{(i,j)}\}$. Instead of selecting a limited number N_{t+1} of surviving

samples through a prediction step, it rather considers a much larger pool ($N_t \times N_{r_{max}}$, with $N_{r_{max}}$ the number of discrete locations in the maximal sphere considered). The dynamic prior is applied retrospectively to these candidates. Another positive side effect of our geometric model and sampling scheme is the elimination of the need for a geometric step parameter. Our filter progresses adaptively, according to the sphere connectivity constraint.

E.3.1 Likelihood Estimation: $p(Y|x_{t+1}^{(i,j)})$

Our likelihood function is a conjunctive combination of three terms:

$$p(Y|x_{t+1}^{(i,j)}) = F(r_t^{(i,j)})E(r_t^{(i,j)})G(x_{t+1}^{(i,j)}) \tag{E.1}$$

The term $F(r_t^{(i,j)})$ measures the gradient flow through the surface of the sphere $S_t^{(i,j)}$ of radius $r_t^{(i,j)} = |x_{t+1}^{(i,j)} - x_t^{(i)}|$ centered on $x_t^{(i)}$:

$$F(r_t^{(i,j)}) = \oint_{S_t^{(i,j)}} \langle \nabla\vec{(I)}, \vec{n} \rangle dS_t^{(i,j)}$$

where $\nabla\vec{(I)}$ is the image gradient and $\vec{n}$ the unit normal to $S_t^{(i,j)}$. It quantifies the alignment of the sphere surface with the image gradient vector field. Similar terms can be found in previous works dedicated to the extraction or filtering of elongated structures Vasilevskiy and Siddiqi (2002); Holtzman-Gazit et al. (2006) or in the minimization scheme of geodesic active contours Caselles et al. (1997).

The term $E(r_t^{(i,j)})$ corresponds to the response of the morphological erosion by the sphere $S_t^{(i,j)}$ at point $x_t^{(i)}$:

$$E(r_t^{(i,j)}) = \min_z \left\{ I(z) \;\middle|\; |z - x_t^{(i)}| \leq r_t^{(i,j)} \right\}$$

where $I(z)$ is the image intensity at location z. This region-based term favors bright spheres and becomes especially useful for small vessels, when the computation of the gradient flow over small discrete spheres becomes unreliable.

The term $G(x_{t+1}^{(i,j)})$ is defined as:

$$G(x_{t+1}^{(i,j)}) = I(x_{t+1}^{(i,j)}) - E(r_t^{(i,j)})$$

This difference term counter-balances the shrinkage effect introduced by $E(r_t^{(i,j)})$ by favoring points with high intensity and low erosion response, *i.e.*, intuitively, locations far from $x_t^{(i)}$. We stress that this term depends on the individual location of $x_{t+1}^{(i,j)}$, not just on the sphere it lies on. It discriminates against locations along the border of the vessel and favors hyper-intense locations toward the centerline. The effects of each term are illustrated in Figure E.3.

In addition, we constrained and normalized image intensities within a range of 0 to 300 Hounsfield units (typical intensities of contrast-enhanced blood pool in CT acquisitions), in order to reduce the influence of hyper-intense structures such as calcifications.

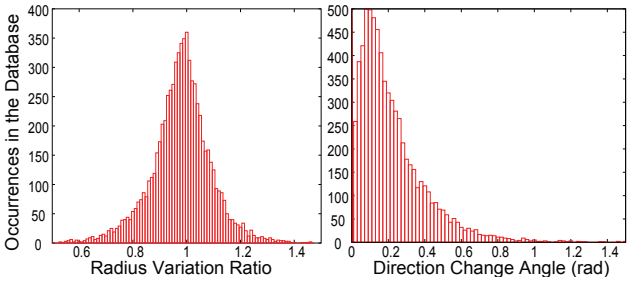
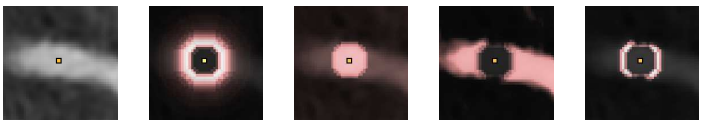


Figure E.2: Learned distributions of the dynamic priors. Left: radius variations. Right: direction variations.



(a) (b) (c) (d) (e)

Figure E.3: Illustration of the likelihood terms. (a) original image and sample position $x_t^{(i)}$; (b) gradient flow term $F(r_t^{(i,j)})$ for all potential successors of $x_t^{(i)}$ overlayed in red; (c) erosion term $E(r_t^{(i,j)})$; (d) "growing" term $G(x_{t+1}^{(i,j)})$; (e) likelihood image, conjunctive combination of the terms (b), (c) and (d).

E.3.2 Dynamic Prior: $p(x_{t+1}^{(i,j)}|x_t^{(i)})$

Vessel dynamics are described in terms of radius and direction variations, considered as independent variables:

$$p(x_{t+1}^{(i,j)}|x_t^{(i)}) = p(r_t^{(i,j)}|r_{t-1}^{(i)})p(\vec{d}_t^{(i,j)}|\vec{d}_{t-1}^{(i)})$$

Both distributions were learned from a pool of 19 manually segmented cardiac CT datasets. Radius variation was evaluated as the ratio $\frac{r_{t+1}}{r_t}$ and direction variation as the angle $\arccos(\langle \frac{\vec{d}_{t+1}}{|\vec{d}_{t+1}|}, \frac{\vec{d}_t}{|\vec{d}_t|} \rangle)$. Samples from the manually segmented database allowed us to build detailed histograms of these distributions (see Figure E.2). These non-parametric forms are used directly to valuate $p(r_t^{(i,j)}|r_{t-1}^{(i)})$ and $p(\vec{d}_t^{(i,j)}|\vec{d}_{t-1}^{(i)})$. ‘

E.4 Evaluation and Results

Our method was evaluated on 19 cardiac CT datasets, for both left and right main coronary branches. Ground-truth segmentations were obtained by detailed manual delineation of the centerlines and associated cross-sections. For each dataset to be tested, dynamic priors were learned on the remaining 18 cases (Leave-One-Out rule) and quantitative results were averaged on 10 runs.

A result example is given in Figure E.4. Besides the sample populations $\{\{x_t^{(i)}, w_t^{(i)}\}_{i=1}^{N_t}\}_{t=0}^{T}$, a key output of the algorithm is the maximum *a posteriori* (MAP) chain $\{x_0^*, ..., x_T^*\}$. It is directly extracted by selecting x_T^* as the sample with highest weight at time-step T and by backtracking its ancestors. From this chain, an approximate *centerline* can be derived by

interpolating a curve linking the spheres centers. Similarly, a volumic representation can be obtained thanks to the associated radius information. We emphasize the conciseness and usability of this MAP result, allowing, for instance, the generation of curved planar reformation views (see Figure E.5(b)).

The algorithm is controlled by a threshold on the length of the extracted branch (9cm for coronaries) and a low threshold on the gradient flow. These constraints were sufficient to prevent "leaking" problems in all our experiments. Remaining parameters are the initial conditions (original seed point at the aortic root), the maximum vessel radius r_{max} (fixed to 6mm for coronaries) and the number of samples selected at each iteration (we used $N_t = 100$, $\forall t$).

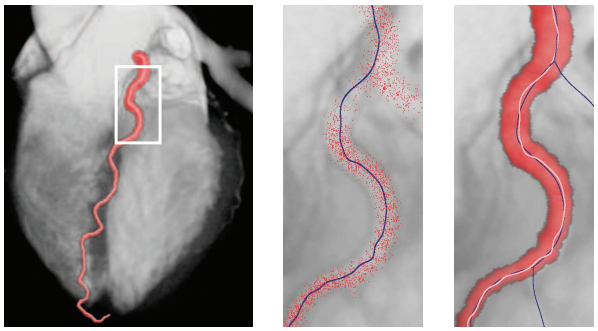
Table E.1 summarizes quantitative results for the 19 datasets. To evaluate the overall robustness of the algorithm, we compared the MAP results to the manual segmentations through the similarity index proposed in Metz et al. (2007): $SI = \frac{2TP}{2TP+FN+FP}$, where TP corresponds to true positive portions (matching segments), FP to false positives and FN to false negatives. We obtained a very high mean similarity factor of $94.8\% \pm 6.2\%$. Early stop was observed in one stenotic case. False positive segments were encountered at the distal end of some branches and directly depend on the stopping criterion.

In terms of accuracy, results for centerline positioning (see also Figure E.4(c)) and radius estimation were satisfactorily in the order of the data resolution (intra-slice resolution of 0.33×0.33mm, inter-slice resolution from 0.33 to 0.6 mm). We stress that we focus more on robust tracking than accurate surface segmentation. Our algorithm output can serve as a close initialization for subsequent refinements, simplifying such tasks through the availability of the centerline and radius information.

Additionally, we observed that nearly optimal tracking performance could be obtained with a very limited number of samples ($N_t \simeq 20$). Increasing the number of samples improves the accuracy of the centerline and radius estimations. Nearly maximal accuracy is reached for $N_t = 100$ samples in all cases. Our original approach, which builds the *entire* distribution of potential successors prior to sampling, potentially explains the very low number of samples needed to attain acceptable tracking performance. This compares very favorably with similar particle-based works such as Florin et al. (2005) and Schaap et al. (2007a), which report the use of 1000 and 500 samples, respectively. Results also appeared to be satisfactorily stable to additive Gaussian noise. Until a noise standard of $\sim$ 100 Hounsfield units (H.U.), which can be considered as a reasonably high noise level by cardiac CT standards, similarity and accuracy measures are nearly unaffected. Then, increasing noise first impacts centerline and radius estimation accuracy. With a noise standard deviation superior to 200 H.U., the geometric information of the image is significantly altered and similarity performance deteriorates as the algorithm fails to track thin vessels disconnected by such a high noise level.

Overall, our Bayesian, multi-hypotheses framework ensures robustness by balancing prior knowledge and information extracted from the image. It proved to be robust to local anatomical anomalies such as aneurysms, stenoses and stents, as illustrated in Figure E.5.

Computational efficiency is promising, especially given the high number of hypotheses evaluated by our sampling scheme. On average, the extraction of a main coronary branch is performed in less than 2 minutes on a Pentium 4 3Ghz computer. The computational time is directly proportional to the number of samples, providing the user with some degree of scalability between computational efficiency and centerline and radius accuracy.



(a) (b) (c)

Figure E.4: Result sample. (a) 3D rendering of the heart with the extracted coronary branch; (b) detail of the samples used by the filter, with the maximum *a posteriori* (MAP) chain (dark blue curve); (c) detail of the MAP centerline with the external envelope of the associated spheres, overlayed with the manually segmented centerlines (dark blue curves).

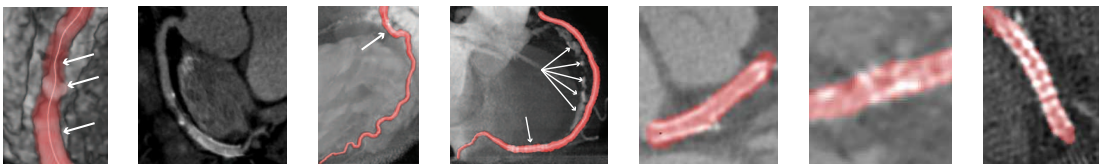
Measure	Mean	Standard Deviation
similarity index	0.948	0.062
centerline position error	0.4mm	0.33mm
radius estimation error	0.33mm	0.21mm

Table E.1: Quantitative evaluation results (see text).

E.5 Conclusion and Perspectives

In this paper, we presented a new tracking algorithm for the segmentation of vascular structures from medical images. Our Bayesian stochastic filter is inspired by recent developments in particle filtering for image segmentation. Novel aspects of our method include a constrained, medial-based geometric model, the use of prior dynamic knowledge learned from a ground-truth database and a new sampling scheme able to cover hypothesis distributions broader than classical particle filters while remaining computationally efficient. Validation of the method for 3D CT coronary artery extraction exhibits very promising results in terms of robustness and accuracy.

We are currently working on a multi-branch version of the algorithm, exploring original solutions to cope with classical problems of loss of diversity linked to importance sampling Doucet et al. (2001). Other refinements include more elaborate dynamics and better stopping criteria. Last but not least, we emphasize the generality of our approach. By simply changing the maximum vessel radius r_{max}, we obtained promising preliminary results on peripheral CT data, as illustrated in Figure E.6. More generally, our method can be directly applied in 2D and the geometric model is not specific to vascular structures. One could imagine, for instance, to adapt our algorithm for road extraction from 2D satellite images.



(a) (b) (c) (d) (e) (f) (g)

Figure E.5: Result samples on data with local anomalies (indicated by arrows). (a) result in presence of a stenosis, an aneurysm and a stent; (b) associated curved planar reformation (CPR) view; (c) result with a stent, high curvature and thin distal end; (d) result with neighboring calcified branch and a stent; (e), (f) and (g) multi-planar reformation views of different results in presence of stents and calcifications.

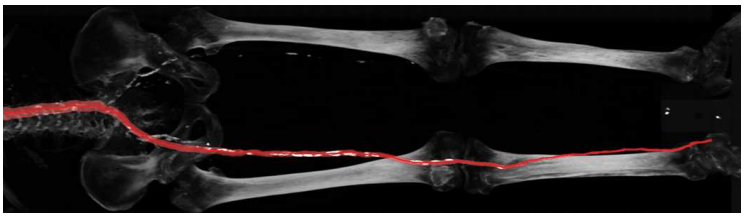


Figure E.6: Preliminary result on peripheral CT data. Segmentation of the descending aorta and right iliac artery.

Bibliography

Adalsteinsson, D., Sethian, J. A., 1995. A fast level set method for propagating interfaces. J. Comput. Phys. 118 (2), 269–277.

Agam, G., Armato III, S. G., Wu, C., 2005. Vessel tree reconstruction in thoracic CT scans with application to nodule detection. IEEE Trans. Med. Imaging 24 (4), 486–499.

Agam, G., Wu, C., 2005. Probabilistic modeling-based vessel enhancement in thoracic CT scans. In: Proc. IEEE Conf. Comput. Vision and Pattern Recognit. pp. 684–689.

Allen, K., Yau, C., Noble, J., 2008. A recursive, stochastic vessel segmentation framework that robustly handles bifurcations. In: Proc. Med. Image Underst. and Analysis.

Angelini, E., Jin, Y., Laine, A., 2005. Handbook of Biomedical Image Analysis- Registration Models. Kluwer Academic/ Plenum Publishers, Ch. State-of-the-Art of Levelset Methods in Segmentation and Registration of Medical Imaging Modalities, pp. 47–102.

Antiga, L., Steinman, D., 2004. Robust and objective decomposition and mapping of bifurcating vessels. IEEE Trans. Med. Imaging 23 (6), 704–713.

Appleton, B., Sun, C., 2003. Circular shortest paths by branch and bound. Pattern Recognit. Letters 36 (11), 2513–2520.

Armande, N., Montesinos, P., Monga, ., 1996. A 3D thin nets extraction method for medical imaging. In: Proc. IEEE Int. Conf. Pattern Recognit. p. 642.

Arulampalam, M., Maskell, S., Gordon, N., Clapp, T., Sci, D., Organ, T., Adelaide, S., 2002. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. IEEE Trans. on Signal Process. 50 (2), 174–188.

Avants, B. B., Williams, J. P., 2000. An adaptive minimal path generation technique for vessel tracking in CTA/CE-MRA volume images. In: Proc. Med. Image Comput. Comput. Assist. Interv. pp. 707–716.

Ayache, N., 1995. Medical computer vision, virtual reality and robotics. Image and Vision Computing 13 (4), 295–313.

Aylward, S., Bullitt, E., 2002. Initialization, noise, singularities, and scale in height ridge traversal for tubular object centerline extraction. IEEE Trans. Med. Imaging 21, 61–75.

Aylward, S., Bullitt, E., Pizer, S., Eberly, D., 1996. Intensity ridge and widths for tubular object segmentation and description. In: Proc. IEEE Mathematical Methods in Biomedical Image Analysis. pp. 131–138.

Azencot, J., Orkisz, M., 2003. Deterministic and stochastic state model of right generalized cylinder (RGC-sm): application in computer phantoms synthesis. Graphical models 65 (6), 323–350.

Bauer, C., Bischof, H., 2008a. Edge based tube detection for coronary artery centerline extraction. The Insight Journal.

Bauer, C., Bischof, H., 2008b. Extracting curve skeletons from gray value images for virtual endoscopy. In: Proc. Medical Imaging and Augmented Reality. pp. 393–402.

Bauer, C., Bischof, H., 2008c. A novel approach for detection of tubular objects and its application to medical image analysis. In: Proc. DAGM Symp. on Pattern Recognit. pp. 163–172.

Behrens, T., Rohr, K., , Stiehl, H., 2001. Segmentation of tubular structures in 3D images using a combination of the hough transform and a kalman filter. In: Proc. DAGM-Symp. Pattern Recognit. Vol. 2191. pp. 406–413.

Benmansour, F., Cohen, L. D., 2009. A new interactive method for coronary arteries segmentation based on tubular anisotropy. In: Proc. IEEE Int. Symp. Biom. Imaging. p. 41.

Benmansour, F., Cohen, L. D., Law, M., Chung, A., 2009. Tubular anisotropy for 2d vessels segmentation. In: Proc. IEEE Conf. Comput. Vision and Pattern Recognit.

Binford, T. O., 1971. Visual perception by computer. In: Proc. IEEE Systems and Control.

Blom, J., 1991. Affine invariant corner detection. Ph.D. thesis, Utrecht University.

Blum, H., 1967. A transformation for extracting new descriptors of shape. Models for the Perception of Speech and Visual Form, 362–380.

Blum, H., 1973. Biological shape and visual science. I. J Theor Biol 38 (2), 205–87.

Boldak, C., Rolland, Y., Toumoulin, C., Coatrieux, J., 2003. An improved model-based vessel tracking algorithm with application to computed tomography angiography. Biocybernetics Biomed. Eng. 23, 41–63.

Boskamp, T., Rinck, D., Link, F., K¸mmerlen, B., Stamm, G., Mildenberger, P., 2004. New vessel analysis tool for morphometric quantification and visualization of vessels in CT and MR imaging data sets. Radiographics 24, 287–297.

Bouix, S., Siddiqi, K., Tannenbaum, A., 2005. Flux driven automatic centerline extraction. Med. Image Anal. 9 (3), 209–221.

Brady, M., Asada, H., 1984. Smoothed local symmetries and their implementation. Int. J. Robotics Research 3 (3), 36.

Broche, B. A., Evans, A. C., Collins, L., 2006. A new improved version of the realistic digital brain phantom. NeuroImage 32 (1), 138–145.

Bruijns, J., 2001. Fully-automatic branch labelling of voxel vessel structures. Proc. Vision Model. and Vis., 341–350.

Buhler, K., Felkel, P., Cruz, A. L., 2002. Geometric methods for vessel visualization and quantification - a survey. Tech. rep. tr vrvis 2002 035, VRVis Research Center, Vienna, Austria.

Bullitt, E., Aylward, S. R., Liu, A., Stone, J., Mukherji, S. K., Coffey, C., Gerig, G., Pizer, S. M., 1999. 3D graph description of the intracerebral vasculature from segmented MRA and tests of accuracy by comparison with x-ray angiograms. In: Proc. Inf. Process. Med. Imaging. pp. 308–321.

Cademartiri, F., La Grutta, L., Palumbo, A., Malagutti, P., Pugliese, F., Meijboom, W., Baks, T., Mollet, N., Bruining, N., Hamers, R., et al., 2007. Non-invasive visualization of coronary atherosclerosis: state-of-art. J. Cardiovasc. Med. 8 (3), 129–137.

Cai, W., Dachille, F., Harris, G., Yoshida, H., 2006. Vesselness propagation: a fast interactive vessel segmentation method. Proc. SPIE Med. Imaging 6144, 1343–1351.

Canny, J., 1986. A computational approach to edge detection. IEEE Trans. Pattern Anal. Mach. Intell. 8 (6), 679–698.

Carpenter, J., Clifford, P., Fearnhead, P., 1999. Improved particle filter for nonlinear problems. In: Proc. IEEE Radar, Sonar and Navigation. Vol. 146. pp. 2–7.

Carrillo, J., Hernàndez Hoyos, M., Dávila, E., Orkisz, M., 2007. Recursive tracking of vascular tree axes in 3D medical images. Int. J. Comput. Assist. Radiol. and Surg. 1 (6), 331–339.

Carrillo, J. F., Orkisz, M., Hernàndez Hoyos, M., 2005. Extraction of 3D vascular tree skeletons based on the analysis of connected components evolution. In: Proc. Comput. Anal. Images and Patterns. pp. 604–611.

Casella, G., Robert, C., 1996. Rao-Blackwellisation of sampling schemes. Biometrika 83 (1), 81–94.

Caselles, V., Kimmel, R., Sapiro, G., 1997. Geodesic active contours. Int. J. Comput. Vision 22 (1), 61–79.

Chan, T., Vese, L., 2001. Active contours without edges. IEEE Trans. Image Process. 10 (2), 266–277.

Cheng, Y., 1995. Mean shift, mode seeking, and clustering. IEEE Trans. Pattern Anal. Mach. Intell. 17 (8), 790–799.

Chopin, N., 2004. Central limit theorem for sequential Monte Carlo methods and its application to Bayesian inference. Annals of statistics, 2385–2411.

Choyke, P., Yim, P., Marcos, H., Ho, V., Mullick, R., Summers, R., 2001. Hepatic MR angiography: a multiobserver comparison of visualization methods. Am. J. Roentgenolog. 176 (2), 465–70.

Chung, A. C. S., Noble, J. A., 1999. Statistical 3D vessel segmentation using a Rician distribution. In: Proc. Med. Image Comput. Comput. Assist. Interv. pp. 82–89.

Chung, A. C. S., Noble, J. A., Summers, P. E., 2004. Vascular segmentation of phase contrast magnetic resonance angiograms based on statistical mixture modeling and local phase coherence. IEEE Trans. Med. Imaging 23 (12), 1490–1507.

Cohen, L., Kimmel, R., 1997. Global minimum for active contour models: A minimal path approach. Int. J. Comput. Vision 24 (1), 57–78.

Cohen, L. D., Deschamps, T., 2001. Grouping connected components using minimal path techniques. In: Proc. IEEE Conf. Comput. Vision and Pattern Recognit. Vol. 2. pp. 102–109.

Comaniciu, D., 2003. An algorithm for data-driven bandwidth selection. IEEE Trans. Pattern Anal. Mach. Intell. 25 (2), 281–288.

Comaniciu, D., Meer, P., 2002. Mean shift: a robust approach toward feature space analysis. IEEE Trans. Pattern Anal. Mach. Intell. 24 (5), 603–619.

Comaniciu, D., Ramesh, V., Meer, P., 2001. The variable bandwidth mean shift and data-driven scale selection. In: Proc. IEEE Int. Conf. Comput. Vision. Vol. 1. pp. 438–335.

Conway, J., Sloane, N., 1999. Sphere packings, lattices and groups. Springer.

Cormen, T., Leiserson, C., Rivest, R., Stein, C., 2001. Introduction to Algorithms. MIT Press.

de Bruijne, M., van Ginneken, B., Niessen, W., Loog, M., Viergever, M., 2003a. Model-based segmentation of abdominal aortic aneurysms in CTA images. In: Proc. SPIE Med. Imaging. Vol. 5032. pp. 1560–1571.

de Bruijne, M., van Ginneken, B., Viergever, M. A., Niessen, W. J., 2003b. Adapting active shape models for 3D segmentation of tubular structures in medical images. In: Proc. Inf. Process. Med. Imaging. pp. 136–147.

Delingette, H., Montagnat, J., 2001. Shape and Topology Constraints on Parametric Active Contours. Comput. Vision and Image Underst. 83 (2), 140–171.

Deriche, R., 1992. Recursively implementing the Gaussian and its derivatives. In: Proc. IEEE Int. Conf. Image Process. pp. 263–267.

Deschamps, T., 2001. Curve and shape extraction with minimal path and level-sets techniques - applications to 3D medical imaging. Ph.D. thesis, Université Paris-IX Dauphine, Place du maréchal de Lattre de Tassigny, 75775 Paris Cedex.

Deschamps, T., Cohen, L., 2000. Minimal paths in 3D images and application to virtual endoscopy. In: Proc. Eur. Conf. Comput. Vision. pp. 543–560.

Deschamps, T., Cohen, L., 2001. Fast extraction of minimal paths in 3D images and applications to virtual endoscopy. Med. Image Anal. 5 (4), 281–299.

Deschamps, T., Cohen, L., 2002a. Fast extraction of tubular and tree 3D surfaces with front propagation methods. In: Proc. IEEE Int. Conf. Pattern Recognit. p. 10731.

Deschamps, T., Cohen, L., 2002b. Geometric Methods in Bio-Medical Image Processing. Mathematics and Visualization. Springer, Ch. Grouping connected components using minimal path techniques.

Deschamps, T., Ebeid, S., Cohen, L., Mar. 1999. Image processing method, system and apparatus for processing an image representing a tubular structure and for constructing a path related to said structure. Patent Pending.

Deschamps, T., Schwartz, P., Trebotich, D., Colella, P., Saloner, D., Malladi, R., 2004. Vessel segmentation and blood flow simulation using level-sets and embedded boundary methods. Int. Congress Series 1268, 75–80.

Descoteaux, M., Collins, D. L., Siddiqi, K., 2004. A multi-scale geometric flow for segmenting vasculature in MRI. In: Proc. Comput. Vision and Math. Methods in Med. and Biomed. Image Anal., ECCV. pp. 169–180.

Dijkstra, E. W., 1959. A note on two problems in connection with graphs. Numerische Mathematik 1 1, 269–271.

Dikici, E., O'Donnell, T., Grady, L., Setser, R., White, R., 2008. Coronary artery centerline tracking using axial symmetries. The Insight Journal.

Douc, R., Cappe, O., Polytech, E., Palaiseau, F., 2005. Comparison of resampling schemes for particle filtering. In: Proc. Int. Symp. Image Signal Process. Anal. pp. 64–69.

Doucet, A., de Freitas, N., Gordon, N., 2001. Sequential Monte Carlo Methods in Practice. Springer-Verlag.

Doucet, A., Godsill, S., Andrieu, C., 2000. On sequential Monte Carlo sampling methods for Bayesian filtering. Statistics and computing 10 (3), 197–208.

Duncan, J., Ayache, N., 2000. Medical image analysis: progress over two decades and the challenges ahead. IEEE Trans. Pattern Anal. Mach. Intell. 22 (1), 85–106.

Eberly, D., Gardner, R., Morse, B., Pizer, S., Scharlach, C., 1994. Ridges for image analysis. J. Math. Imaging and Vision 4 (4), 353–373.

Falcao, A., Udupa, J., Miyazawa, F., 2000. An ultra-fast user-steered image segmentation paradigm: live-wire on the fly. IEEE Trans. Med. Imaging 19 (1), 55–62.

Fanucci, E., Orlacchio, A., Pocek, M., 1988. The vascular geometry of human arterial bifurcations. Invest. Radiol. 23 (10), 713–8.

Felkel, P., Wegenkittl, R., Kanitsar, A., 2001. Vessel Tracking in Peripheral CTA Datasets–An Overview. In: Proc. Spring Conference on Computer Graphics. pp. 269–278.

Flaaris, J., Volden, M., Haase, J., Østergaard, L., 2004. Method for modelling cerebral blood vessels and their bifurcations using circular, homogeneous, generalised cylinders. Med. Biol. Eng. Comput. 42 (2), 171–177.

Flasque, N., Desvignes, M., Constans, J., Revenu, M., 2001. Acquisition, segmentation and tracking of the cerebral vascular tree on 3D magnetic resonance angiography images. Med Image Anal 5 (3), 173–83.

Flórez Valencia, L., Azencot, J., Vincent, F., Orkisz, M., Magnin, I., 2006. Segmentation and quantification of blood vessels in 3d images using a right generalized cylinder state model. In: Proc. IEEE Int. Conf. Image Process. pp. 2441–2444.

Florin, C., Moreau-Gobard, R., Williams, J., 2004. Automatic heart peripheral vessels segmentation based on a normal mip ray casting technique. In: Proc. Med. Image Comput. Comput. Assist. Interv. pp. 483–490.

Florin, C., Paragios, N., Williams, J., 2005. Particle filters, a quasi-Monte Carlo solution for segmentation of coronaries. In: Proc. Med. Image Comput. Comput. Assist. Interv. pp. 246–253.

Florin, C., Paragios, N., Williams, J., 2006. Globally optimal active contours, sequential Monte-Carlo and on-line learning for vessel segmentation. In: Proc. Eur. Conf. Comput. Vision. pp. 476–489.

Fox, D., 2003. Adapting the sample size in particle filters through KLD-sampling. Int. J. Robotics Research 22 (12), 985–1003.

Frangi, A., Niessen, W., Hoogeveen, R., van Walsum, T., Viergever, M., 1999a. Model-based quantitation of 3-D magnetic resonance angiographic images. IEEE Trans. Med. Imaging 18 (10), 946–956.

Frangi, A., Niessen, W., Nederkoorn, P., Elgersma, O., Viergever, M., 2000. Three-dimensional model-based stenosis quantification of the carotid arteries from contrast-enhance MR angiography. In: Proc. IEEE Mathematical Methods in Biomedical Image Analysis. pp. 110–118.

Frangi, A. F., Niessen, W. J., Hoogeveen, R. M., van Walsum, T., Viergever, M. A., 1999b. Quantitation of vessel morphology from 3D MRA. In: Proc. Med. Image Comput. Comput. Assist. Interv. pp. 358–367.

Frangi, A. F., Niessen, W. J., Vincken, K. L., Viergever, M. A., 1998. Multiscale vessel enhancement filtering. In: Proc. Med. Image Comput. Comput. Assist. Interv. Vol. 1496. pp. 130–137.

Freedman, D., Kisilev, P., 2009. Fast mean shift by compact density representation. In: Proc. IEEE Conf. Comput. Vision and Pattern Recognit.

Fridman, Y., 2004. Extracting Branching Object Geometry via Cores. Ph.D. thesis, University of North Carolina.

Fridman, Y., Pizer, S. M., Aylward, S. R., Bullitt, E., 2003. Segmenting 3D branching tubular structures using cores. In: Proc. Med. Image Comput. Comput. Assist. Interv. pp. 570–577.

Friman, O., Hindennach, M., Peitgen, H.-O., 2008a. Template-based multiple hypotheses tracking of small vessels. In: Proc. IEEE Int. Symp. Biom. Imaging. pp. 1047–1050.

Friman, O., K¸hnel, C., Peitgen, H., 2008b. Coronary artery centerline extraction using multiple hypothesis tracking and minimal paths. The Insight Journal.

Fukunaga, K., 1990. Introduction to statistical pattern recognition, second ed. Academic Press.

Fukunaga, K., Hostetler, L., 1975. The estimation of the gradient of a density function, with applications in pattern recognition. IEEE Trans. Information Theory 21 (1), 32–40.

Funka-Lea, G., Boykov, Y., Florin, C., Jolly, M., Moreau-Gobard, R., Ramaraj, R., Rinck, D., 2006. Automatic heart isolation for CT coronary visualization using graph-cuts. In: Proc. IEEE Int. Symp. Biom. Imaging. pp. 614–617.

Furst, J., 1999. Height ridges of oriented medialness. Ph.D. thesis, Department of Computer Science, University of North Carolina at Chapel Hill.

Gan, R., Chung, A. C. S., Wong, W. C. K., Yu, S. C. H., 2004. Vascular segmentation in three-dimensional rotational angiography based on maximum intensity projections. In: Proc. IEEE Int. Symp. Biom. Imaging. pp. 133–136.

Geiger, D., Gupta, A., Costa, L., Vlontzos, J., 1995. Dynamic programming for detecting, tracking, and matching deformable contours. IEEE Trans. Pattern Anal. Mach. Intell. 17 (3), 294–302.

Geman, D., Jedynak, B., 1996. An active testing model for tracking roads in satellite images. IEEE Trans. Pattern Anal. Mach. Intell. 18 (1), 1–14.

Georgescu, B., Shimshoni, I., Meer, P., 2003. Mean shift based clustering in high dimensions: A texture classification example. In: Proc. IEEE Int. Conf. Comput. Vision. pp. 456–463.

Gerig, G., Kubler, O., Kikinis, R., Jolesz, F., 1992. Nonlinear anisotropic filtering of MRI data. IEEE Trans. Med. Imaging 11 (2), 221–232.

Geweke, J., 1989. Bayesian inference in econometric models using Monte Carlo integration. Econometrica: Journal of the Econometric Society, 1317–1339.

Gilks, W., Berzuini, C., 2001. Following a moving target-Monte Carlo inference for dynamic Bayesian models. J. R. Statist. Soc. B, 127–146.

Gilks, W., Spiegelhalter, D., 1996. Markov chain Monte Carlo in practice. Chapman & Hall/CRC.

Godsill, S., Clapp, T., 2001. Improvement strategies for Monte Carlo particle filters. In: Sequential Monte Carlo Methods in Practice. Springer-Verlag, pp. 139–158.

Godsill, S., Doucet, A., West, M., 2001. Maximum a posteriori sequence estimation using Monte Carlo particle filters. Annals of the Institute of Statistical Mathematics 53 (1), 82–96.

Gong, R., Wörz, S., Rohr, K., 2003. Segmentation of Coronary Arteries of the Human Heart from 3D Medical Images. In: Proc. Workshop Bildverarbeitung fur die Medizin. pp. 66–70.

Gordon, N., Salmond, D., Smith, A., 1993. Novel approach to nonlinear/non-Gaussian Bayesian state estimation. In: Proc. Inst. Elect. Eng. Vol. 140. pp. 107–113.

Grady, L., 2006. Fast, Quality, Segmentation of Large Volumes-Isoperimetric Distance Trees. ECCV 3953, 449–462.

Grady, L., Schwartz, E., 2006. Isoperimetric graph partitioning for image segmentation. IEEE Trans. Pattern Anal. Mach. Intell. 28 (3), 469–475.

Gulsun, M. A., Tek, H., 2006. 3d construction of coronary arteries. In: Int. Workshop on Comput. Vision for Intravascular and Intracardiac Imaging, Proc. Med. Image Comput. Comput. Assist. Interv.

Gulsun, M. A., Tek, H., 2008a. Geometric modeling of tubular structures. In: Proc. IEEE Conf. Comput. Vision and Pattern Recognit.

Gulsun, M. A., Tek, H., 2008b. Robust vessel tree modeling. In: Proc. Med. Image Comput. Comput. Assist. Interv.

Hardin, R. H., Sloane, N. J. A., Smith, W. D., 2000. Tables of spherical codes with icosahedral symmetry. http://www.research.att.com/~njas/icosahedral.codes/.

Hassouna, M. S., Farag, A. A., Hushek, S. G., Moriarty, T., 2003. Statistical-based approach for extracting 3D blood vessels from tof-myra data. In: Proc. Med. Image Comput. Comput. Assist. Interv. pp. 680–687.

Hernandez, M., Frangi, A., Sapiro, G., 2003. Three-dimensional segmentation of brain aneurysms in CTA using non-parametric region-based information and implicit deformable models: Method and evaluation. In: Proc. Med. Image Comput. Comput. Assist. Interv. pp. 594–602.

Hernàndez Hoyos, M., 2002. Segmentation anisotrope 3D pour la quantification en imagerie vasculaire par rÈsonance magnÈtique (in french). Ph.D. thesis, École Doctorale des Sciences de l'IngÈnieur de Lyon, CREATIS UMR-CNRS 5515, Lyon, France.

Hernàndez Hoyos, M., Anwander, A., Orkisz, M., Roux, J.-P., Douek, P., Magnin, I. E., 2000. A deformable vessel model with single point initialization for segmentation, quantification and visualization of blood vessels in 3D MRA. In: Proc. Med. Image Comput. Comput. Assist. Interv. pp. 735–745.

Hernàndez Hoyos, M., Orkisz, M., Douek, P., Magnin, I., 2005. Assessment of carotid artery stenoses in 3D contrast-enhanced magnetic resonance angiography, based on improved generation of the centerline. Int. J. Mach. Graph. & Vis. 14 (4), 378.

Hernàndez Hoyos, M., Orłowski, P., Piatkowska-Janko, E., Bogorodzki, P., Orkisz, M., 2006. Vascular centerline extraction in 3D MR angiograms for phase contrast MRI blood flow measurement. Int. J. Comput. Assist. Radiol. and Surg. 1 (1), 51–61.

Higuchi, T., 1997. Monte Carlo filter using the genetic algorithm operators. J. Statist. Comput. Simul. 59 (1), 1–23.

Holtzman-Gazit, M., Kimmel, R., Peled, N., Goldsher, D., 2006. Segmentation of thin structures in volumetric medical images. IEEE Trans. Image Process. 15 (2), 354–363.

Hoogeveen, R., Bakker, C., Viergever, M., 1998. Limits to the accuracy of vessel diameter measurement in MR angiography. J. Magn. Reson. Imaging 8 (6).

Huang, Q., Stockman, G. C., 1993. Generalized tube model: recognizing 3D elongated objects from 2D intensity images. In: Proc. IEEE Conf. Comput. Vision and Pattern Recognit. pp. 104–109.

Hürzeler, M., Künsch, H., 1998. Monte Carlo approximations for general state-space models. J. Computational and graphical statistics, 175–193.

Isard, M., Blake, A., 1998a. A mixed-state condensation tracker with automatic model-switching. In: Proc. IEEE Int. Conf. Comput. Vision. pp. 107–112.

Isard, M., Blake, A., 1998b. Condensation: conditional density propagation for visual tracking. Int. J. Comput. Vision 29, 5–28.

Jomier, J., LeDigarcher, V., Aylward, S. R., 2005. Automatic vascular tree formation using the mahalanobis distance. In: Proc. Med. Image Comput. Comput. Assist. Interv. pp. 806–812.

Kalman, R., 1960. A new approach to linear filtering and prediction problems. Basic Eng. 82 (1), 35–45.

Kanitsar, A., Fleischmann, D., Wegenkittl, R., Felkel, P., Groller, E., 2002. CPR-curved planar reformation. IEEE Visualization, 37–44.

Kass, M., Witkin, A., Terzopoulos, D., 1987. Snakes: Active contour models. Int. J. Comput. Vision 1, 321–331.

Kawata, Y., Niki, N., Kumazaki, T., 1995. An approach for detecting blood vessel diseases from cone-beam CT image. In: Proc. IEEE Int. Conf. Image Process. Vol. 2. pp. 500–503.

Kawata, Y., Niki, N., Kumazaki, T., 1996. Measurement of blood vessel characteristics for disease detection based on cone-beam CT images. IEEE Trans. Nuclear Science 43 (6), 3348–3354.

Kimmel, R., Bruckstein, A. M., 2003. Regularized laplacian zero crossings as optimal edge integrators. Int. J. Comput. Vision 53 (3), 225–243.

Kimmel, R., Sethian, J., 2001. Optimal algorithm for shape from shading and path planning. J. M. Imaging and Vision 14 (3), 237–244.

Kirbas, C., Quek, F., 2002. 3D wave propagation and traceback in vascular extraction. In: Proc. IEEE Eng. Med. Biol. Soc. Vol. 2. pp. 1078–1079.

Kirbas, C., Quek, F., 2004. A review of vessel extraction techniques and algorithms. ACM Computing Surveys 36 (2), 81–121.

Kirbas, C., Quek, F. K. H., 2003. Vessel extraction in medical images by 3D wave propagation and traceback. In: Proc. IEEE Symp. on BioInf. and BioEng. pp. 174–181.

Kitagawa, G., 1996. Monte Carlo filter and smoother for non-Gaussian nonlinear state space models. J. Computational and graphical statistics, 1–25.

Kitagawa, G., Gersch, W., 1996. Smoothness Priors Analysis of Time Series. Springer.

Kitslaar, P., Frenay, M., Oost, E., Dijkstra, J., B., S., Reiber, J., 2008. Connected component and morpholgy based extraction of arterial centerlines of the heart (cocomobeach). The Insight Journal.

Klaas, M., Briers, M., de Freitas, N., Doucet, A., Maskell, S., Lang, D., 2006. Fast particle smoothing: If i had a million particles. In: Int. Conf. Machine learning. pp. 481–488.

Koller, D., Fratkina, R., 1998. Using learning for approximation in stochastic processes. In: Proc. Int. Conf. Machine Learning. pp. 287–295.

Koller, T., Gerig, G., Székely, G., Dettwiler, D., 1995. Multiscale detection of curvilinear structures in 2D and 3D image data. In: Proc. IEEE Int. Conf. Comput. Vision. pp. 864–869.

Kong, A., Liu, J., Wong, W., 1994. Sequential imputations and Bayesian missing data problems. J. Amer. Statist. Assoc. 89, 278–288.

Konishi, S., Yuille, A., Coughlan, J., Zhu, S., 2003. Statistical edge detection: Learning and evaluating edge cues. IEEE Trans. Pattern Anal. Mach. Intell. 25 (1), 57–74.

Krissian, K., 2002. Flux-based anisotropic diffusion applied to enhancement of 3-D angiogram. IEEE Trans. Med. Imaging 21 (11), 1441.

Krissian, K., Malandain, G., Ayache, N., July 1997. Directional anisotropic diffusion applied to segmentation of vessels in 3D images. In: Scale-Space Theory in Computer Vision (Scale-Space'97). Vol. 1252 of LNCS. Springer, Utrecht, The Netherlands, pp. 345–348.

Krissian, K., Malandain, G., Ayache, N., Vaillant, R., Trousset, Y., 2000. Model-based detection of tubular structures in 3D images. Comput. Vision and Image Underst. 80 (2), 130–171.

La Cruz, A., Straka, M., Kochl, A., Groller, E., Fleischmann, D., 2004. Non-linear model fitting to parameterize diseased blood vessels. In: Proc. IEEE Visu. pp. 393–400.

Lacoste, C., Descombes, X., Zerubia, J., 2005. Point processes for unsupervised line network extraction in remote sensing. IEEE Trans. Pattern Anal. Mach. Intell. 27 (9), 1568–1579.

Lacoste, C., Finet, G., Magnin, I. E., 2006. Coronary tree extraction from x-ray angiograms using marked point processes. In: Proc. IEEE Int. Symp. Biom. Imaging. pp. 157–160.

Lanz, O., 2007. An information theoretic rule for sample size adaptation in particle filtering. In: Proc. Int. Conf. Image Anal. and Process. pp. 317–322.

Larralde, A., Boldak, C., Garreau, M., Toumoulin, C., Boulmier, D., Rolland, Y., 2003. Evaluation of a 3D segmentation software for the coronary characterization in multislice computed tomography. In: Proc. Functional Imaging and Modeling of the Heart. pp. 39–51.

Law, M., Chung, A., 2008. Three dimensional curvilinear structure detection using optimally oriented flux. In: Eur. Conf. Comput. Vision. pp. 368–382.

Law, M., Chung, A., 2009. Efficient implementation for spherical flux computation and its application to vascular segmentation. IEEE Trans. Image Process. 18 (3), 596–612.

Law, T. Y., Heng, P.-A., 2000. Automatic centerline extraction for 3D virtual bronchoscopy. In: Proc. Med. Image Comput. Comput. Assist. Interv. pp. 786–795.

Lee, J., Beighley, P., Ritman, E., Smith, N., 2007. Automatic segmentation of 3D micro-CT coronary vascular images. Med. Image Anal. 11 (6), 630–647.

Lei, T., Udupa, J. K., Saha, P. K., Odhner, D., 2001. Artery-vein separation via MRA - an image processing approach. IEEE Trans. Med. Imaging 20 (8), 689–703.

Lesage, D., Angelini, A., Bloch, I., Funka-Lea, G., 2009a. A review of 3d vessel lumen segmentation techniques: Models, features and extraction schemes. Med. Image Anal. (in press).

Lesage, D., Angelini, E., Bloch, I., Funka-Lea, G., 2008. Medial-based bayesian tracking for vascular segmentation: Application to coronary arteries in 3d CT angiography. In: Proc. IEEE Int. Symp. Biom. Imaging. pp. 268–271.

Lesage, D., Angelini, E., Bloch, I., Funka-Lea, G., 2009b. Bayesian maximal paths for coronary artery segmentation from 3d ct angiograms. In: Proc. Med. Image Comput. Comput. Assist. Interv. (in press).

Lesage, D., Angelini, E., Bloch, I., Funka-Lea, G., 2009c. Design and study of flux-based features for 3d vascular tracking. In: Proc. IEEE Int. Symp. Biom. Imaging. pp. 286–289.

Li, H., Yezzi, A., 2006. Vessels as 4D curves: Global minimal 4D paths to extract 3D tubular surfaces. In: Proc. IEEE Conf. Comput. Vision and Pattern Recognit. p. 82.

Li, H., Yezzi, A., 2007. Vessels as 4-D curves: Global minimal 4-D paths to extract 3-D tubular surfaces and centerlines. IEEE Trans. Med. Imaging 26, 1213–1223.

Li, Q., Sone, S., Doi, K., 2003. Selective enhancement filters for nodules, vessels, and airway walls in two-and three-dimensional CT scans. Med. Physics 30, 2040.

Li, R., Ourselin, S., 2003. Accurate curvilinear modeling for precise measurements of tubular structures. In: Proc. Digital Image Computing: Techniques and Applications. pp. 243–252.

Lin, Q., May 2003. Enhancement, Detection, and Visualization of 3D Volume Data. Ph.D. thesis, Dept. EE, Linköping University, SE-581 83 Linköping, Sweden, dissertations No. 824.

Lindeberg, T., 1994. Scale-Space Theory in Computer Vision. Kluwer Academic Publishers.

Liu, J., Chen, R., 1998. Sequential Monte Carlo methods for dynamic systems. J. Amer. Statist. Assoc. 93, 1032–1044.

Lorenz, C., Carlsen, I., Buzug, T., Fassnacht, C., Weese, J., 1997. Multiscale line segmentation with automatic estimation of width, contrast and tangential direction in 2D and 3D medical images. In: Proc. CVRMed and MRCAS. pp. 233–242.

Lorenz, C., Renisch, S., Schlathoelter, T., Buelow, T., 2003. Simultaneous segmentation and tree reconstruction of the coronary arteries in MSCT images. In: Proc. SPIE Med. Imaging. pp. 167–177.

Lorigo, L. M., Faugeras, O. D., Grimson, W. E. L., Keriven, R., Kikinis, R., Nabavi, A., Westin, C.-F., 2001. Curves: Curve evolution for vessel segmentation. Med. Image Anal. 5, 195–206.

Luboz, V., Wu, X., Krissian, K., Westin, C.-F., Kikinis, R., Cotin, S., Dawson, S., 2005. A segmentation and reconstruction technique for 3d vascular structures. In: Proc. Med. Image Comput. Comput. Assist. Interv. pp. 43–50.

Malandain, G., Bertrand, G., Ayache, N., 1993. Toplogical segmentation of discrete surfaces. In: Int. J. Comput. Vision. Vol. 2. pp. 183–197.

Manniesing, R., Niessen, W., 2004. Local speed functions in level set based vessel segmentation. In: Proc. Med. Image Comput. Comput. Assist. Interv. Vol. 3216. pp. 475–482.

Manniesing, R., Viergever, M., Niessen, W., 2006. Vessel enhancing diffusion A scale space representation of vessel structures. Med. Image Anal. 10 (6), 815–825.

Masutani, Y., Schiemann, T., Höhne, K. H., 1998. Vascular shape segmentation and structure extraction using a shape-based region-growing model. In: Proc. Med. Image Comput. Comput. Assist. Interv. pp. 1242–1249.

Matej, S., Lewitt, R., 1995. Efficient 3D grids for image reconstruction using spherically-symmetric volume elements. IEEE Trans. Nuclear Science 42 (4), 1361–1370.

McInerney, T., Terzopoulos, D., 1996. Deformable models in medical image analysis: A survey. Med. Image Anal. 1, 91–108.

McInerney, T., Terzopoulos, D., 1999. Topology adaptive deformable surfaces for medical image volume segmentation. IEEE Trans. Med. Imaging 18, 840–850.

McInerney, T., Terzopoulos, D., 2000. T-snakes: Topology adaptive snakes. Med. Image Anal. 4 (2), 73–91.

Meijering, E., Niessen, W., Weickert, J., Viergever, M., 2001. Evaluation of Diffusion Techniques for Improved Vessel Visualization and Quantification in Three-Dimensional Rotational Angiography. In: Proc. Med. Image Comput. Comput. Assist. Interv. pp. 177–185.

Metz, C., Schaap, M., Van Der Giessen, A., Van Walsum, T., Niessen, W., 2007. Semi-automatic coronary artery centerline extraction in computed tomography angiography data. In: Proc. IEEE Int. Symp. Biom. Imaging. pp. 856–859.

Metz, C., Schaap, M., van Walsum, T., Niessen, W., 2008a. Two point minimum cost path approach for CTA coronary centerline extraction. The Insight Journal.

Metz, C., Schaap, M., Van Walsum, T., Van der Giessen, A., Weustink, A., Mollet, N., Krestin, G., Niessen, W., 2008b. Rotterdam Coronary Artery Algorithm Evaluation Framework . http://coronary.bigr.nl/.

Mille, J., Boné, R., Cohen, L., 2008. Region-based 2D deformable generalized cylinder for narrow structures segmentation. In: Proc. Eur. Conf. Comput. Vision. Springer, pp. 392–404.

Mille, J., Cohen, L., 2009. Deformable tree models for 2D and 3D branching structures extraction. In: Proc. IEEE Math. Methods Biomed. Image Anal.

Montagnat, J., 1999. Deformable modelling for 3D and 4D medical image segmentation. Ph.D. thesis, Nice - Sophia Antipolis University, Nice, France.

Mowatt, G., Cummins, E., Waugh, N., Walker, S., Cook, J., Jia, X., Hillis, G., Fraser, C., 2008. Systematic review of the clinical effectiveness and cost-effectiveness of 64-slice or higher computed tomography angiography as an alternative to invasive coronary angiography in the investigation of coronary artery disease. Health Technol. Assess. 12 (17).

Murray, C., 1926. The physiological principle of minimum work. i. the vascular system and the cost of blood volume. Proc. National Academy of Sciences of the United States of America 12 (3), 207–214.

Musso, C., Oudjane, N., LeGland, F., 2001. Improving regularised particle filters. In: Sequential Monte Carlo Methods in Practice. Springer-Verlag, pp. 247–271.

Naf, M., Kubler, O., Kikinis, R., Shenton, M., Szekely, G., 1996. Characterization and recognition of 3D organ shape in medical imageanalysis using skeletonization. In: Proc. IEEE Math. Methods Biomed. Image Anal. pp. 139–150.

Nain, D., Yezzi, A. J., Turk, G., 2004. Vessel segmentation using a shape driven flow. In: Proc. Med. Image Comput. Comput. Assist. Interv. pp. 51–59.

Niessen, W., van Swijndregt, A., Elsman, B., Wink, O., Viergever, M., Mali, W., 1999. Enhanced artery visualization in blood pool MRA: Results in the peripheral vasculature. In: Proc. Inf. Process. Med. Imaging. Vol. 1613. pp. 340–345.

O'Donnell, T., Boult, T., Fang, X., Gupta, A., 1994. The extruded generalized cylinder: A deformable model for object recovery. In: Proc. IEEE Conf. Comput. Vision and Pattern Recognit. pp. 174–181.

O'Donnell, T., Dubuisson-Jolly, M.-P., Gupta, A., 1998. A cooperative framework for segmentation using 2D active contours and 3D hybrid models as applied to branching cylindrical structures. In: Proc. IEEE Int. Conf. Comput. Vision. pp. 454–459.

Oka, S., Nakai, M., 1987. Optimality principle in vascular bifurcation. Biorheology 24 (6), 737–51.

Olabarriaga, S., Breeuwer, M., Niessen, W., 2003. Minimum Cost Path Algorithm for Coronary Artery Central Axis Tracking in CT Images. In: Proc. Med. Image Comput. Comput. Assist. Interv. pp. 687–694.

Orkisz, M., Bresson, C., Magnin, I., Champin, O., Douek, P., 1997. Improved vessel visualization in MR angiography by nonlinear anisotropic filtering. Magn. Reson. Med. 37 (6), 914–919.

Orkisz, M., Hernàndez Hoyos, M., Douek, P., Magnin, I., 2000. Advances of blood vessel morphology analysis in 3D magnetic resonance images. Int. J. Mach. Graph. & Vis. 9, 463–471.

Oudjane, N., Musso, C., 2000. Progressive correction for regularized particle filters. In: Proc. Int. Conf. Information Fusion. Vol. 2.

Paladini, G., Azar, F., 2009. An extensible imaging platform for optical imaging applications. In: Proc. SPIE. Vol. 7171. pp. 8–10.

Palagyi, K., Kuba, A., 1998. A 3D 6-subiteration thinning algorithm for extracting medial lines. In: Pattern Recognit. Letters. Vol. 19. pp. 613–627.

Passat, N., Ronse, C., Baruthio, J., Armspach, J.-P., Maillot, C., 2006. Magnetic resonance angiography: From anatomical knowledge modeling to vessel segmentation. Med. Image Anal. 10 (2), 259–274.

Pearl, J., 1984. Heuristics: Intelligent Search Strategies for Computer Problem Solving. Addison-Wesley, Reading, Mass.

Pechaud, M., Keriven, R., G., P., 2009. Extraction of tubular structures over an orientation domain. In: Proc. IEEE Conf. Comput. Vision and Pattern Recognit.

Pitt, M., Shephard, N., 1999. Filtering Via Simulation: Auxiliary Particle Filters. J. Amer. Statist. Assoc. 94 (446), 590–591.

Pizer, S., Eberly, D., Morse, B., Fritsch, D., 1998. Zoom-invariant vision of figural shape: The mathematics of cores. Comput. Vision and Image Underst. 69 (1), 055–071.

Pizer, S., Siddiqi, K., Székely, G., Damon, J., Zucker, S., 2003. Multiscale medial loci and their properties. Int. J. Comput. Vision 55 (2), 155–179.

Pock, T., Janko, C., Beichel, R., Bischof, H., 2005. Multiscale medialness for robust segmentation of 3D tubular structures. In: Proc. Comput. Vision Winter Workshop. pp. 93–102.

Poli, R., Valli, G., 1996. An algorithm for real-time vessel enhancement and detection. Computer Methods and Programs in Biomedicine 52, 1–2.

Ponce, J., Chelberg, D. M., Mann, W. B., 1989. Invariant properties of straight homogeneous generalized cylinders and their contours. IEEE Trans. Pattern Anal. Mach. Intell. 11 (9), 951–966.

Prinet, V., Monaga, O., Ge, C., Xie, S. L., Ma, S. D., 1996. Thin network extraction in 3D images: Application to medical angiograms. In: Proc. IEEE Int. Conf. Pattern Recognit. pp. 386–390.

Pudney, C., 1998. Distance-ordered homotopic thinning: A skeletonization algorithm for 3D digital images. Comput. Vision and Image Underst. 72 (3), 404–413.

Puig, A., Tost, D., Navazo, I., 1997. An interactive cerebral blood vessel exploration system. In: Proc. IEEE Visualization. pp. 433–436.

Quek, F. K. H., Kirbas, C., 2001. Vessel extraction in medical images by wave propagation and traceback. IEEE Trans. Med. Imaging 20 (2), 117–131.

Reinhardt, J., D'Souza, N., Hoffman, E., 1997. Accurate measurement of intrathoracic airways. IEEE Trans. Med. Imaging 16 (6), 820–7.

Reuze, P., Coatrieux, J., Luo, L., Dillenseger, J., 1993. A 3D moment based approach for blood vessel detection and quantification in MRA. Technol. Health Care 1, 181–188.

Rochery, M., Jermyn, I., Zerubia, J., 2005a. New higher-order active contour energies for network extraction. In: Proc. IEEE Int. Conf. Image Process. Vol. 2. pp. 822–825.

Rochery, M., Jermyn, I., Zerubia, J., 2005b. Phase field models and higher-order active contours. Proc. IEEE Int. Conf. Comput. Vision 2, 970–976.

Rochery, M., Jermyn, I., Zerubia, J., 2006. Higher order active contours. Int. J. Comput. Vision 69 (1), 27–42.

Rohr, K., Wörz, S., 2006. High-precision localization and quantification of 3D tubular structures. In: Proc. IEEE Int. Symp. Biom. Imaging. pp. 1160–1163.

Rosamond, W., Flegal, K., Furie, K., Go, A., Greenlund, K., Haase, N., Hailpern, S., Ho, M., Howard, V., Kissela, B., et al., 2008. Heart disease and stroke statistics–2008 update: a report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. Circulation 117 (4), 25–146.

Rouchdy, Y., Cohen, L., 2008. Image segmentation by geodesic voting. Application to the extraction of tree structures from confocal microscope images. In: Proc. IEEE Int. Conf. Pattern Recognit.

Rousson, M., Bai, Y., Xu, C., Sauer, F., 2006. Probabilistic minimal path for automated esophagus segmentation. In: Proc SPIE. Vol. 2. pp. 6144–6153.

Rueckert, D., Burger, P., Forbat, S. M., Mohiaddin, R., Yang, G.-Z., 1997. Automatic tracking of the aorta in cardiovascular MR images using deformable models. IEEE Trans. Med. Imaging 16 (5), 581–590.

Saba, O., Hoffman, E., Reinhardt, J., 2003. Maximizing quantitative accuracy of lung airway lumen and wall measures obtained from x-ray ct imaging. J. Appl. Physiol. 95 (3), 1063–1075.

Saff, E., Kuijlaars, A., 1997. Distributing many points on a sphere. Mathematical Intelligencer 19 (1), 5–11.

Sato, Y., Araki, T., Hanayama, M., Naito, H., Tamura, S., 1998a. A viewpoint determination system for stenosis diagnosis and quantification in coronary angiographic image acquisition. IEEE Trans. Med. Imaging 17 (1), 121–137.

Sato, Y., Nakajima, S., Atsumi, H., Koller, T., Gerig, G., Yoshida, S., Kikinis, R., 1998b. 3D multi-scale line filter for segmentation and visualization of curvilinear structures in medical images. Med. Image Anal. 2, 143–168.

Schaap, M., Manniesing, R., Smal, I., van Walsum, T., van der Lugt, A., Niessen, W. J., 2007a. Bayesian tracking of tubular structures and its application to carotid arteries in cta. In: Proc. Med. Image Comput. Comput. Assist. Interv. pp. 562–570.

Schaap, M., Metz, C., van Walsum, T., van der Giessen, A., Weustink, A., Mollet, N., Bauer, C., Bogunovic', H., Castro, C., Deng, X., Dikici, E., O'Donnell, T., Frenay,

M., Friman, O., Hernàndez Hoyos, M., Kitslaar, P., Krissian, K., K¸hnel, C., Luengo-Oroz, M. A., Orkisz, M., Smedby, O., Styner, M., Szymczak, A., Tek, H., Wang, C., Warfield, S. K., Zambal, S., Zhang, Y., Krestin, G. P., Niessen, W., 2009a. Standardized evaluation methodology and reference database for evaluating coronary artery centerline extraction algorithms. Medical Image Analysis, 701–714.

Schaap, M., Neefjes, L., Metz, C., van der Giessen, A., Weustink, A., Mollet, N., Wentzel, J., van Walsum, T., Niessen, W., 2009b. Coronary lumen segmentation using graph cuts and robust kernel regression. In: Jerry L. Prince, Dzung L. Pham, K. J. M. (Ed.), Proc. Information Process. Medical Imaging. pp. 528–539.

Schaap, M., Smal, I., Metz, C., van Walsum, T., Niessen, W., 2007b. Bayesian tracking of elongated structures in 3D images. In: Proc. Inf. Process. Med. Imaging. Vol. 4584. Springer, pp. 74–85.

Sebbe, R., Gosselin, B., Coche, E., Macq, B., 2003. Pulmonary arteries segmentation and feature extraction through slice marching. In: Proc. ProRISC workshop on Circuits, Systems and Signal Process.

Sekiguchi, H., Sugimoto, N., Eiho, S., Hanakawa, T., Urayama, S., 2005. Blood vessel segmentation for head MRA using branch-based region growing. Systems and Computers in Japan 36 (5), 80–88.

Serra, J., 1983. Image Analysis and Mathematical Morphology. Academic Press, Inc. Orlando, FL, USA.

Serra, J., Salembier, P., 1993. Connected operators and pyramids. Proc. SPIE Image Algebra and Morphol. Image Process. 2030, 65.

Sethian, J. A., February 1996a. A fast marching level set method for monotonically advancing fronts. Proc. National Academy of Sciences of the USA 93 (4), 1591–1595.

Sethian, J. A., 1996b. Level set methods: Evolving interfaces in geometry, fluid mechanics, computer vision, and materials science. Cambridge University Press.

Sethian, J. A., 1999. Level Set Methods and Fast Marching Methods. Cambridge University Press.

Shi, Y., Karl, W., 2005. A fast level set method without solving pdes. In: IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005. Proceedings. (ICASSP '05). Vol. 2. pp. 97–100.

Shikata, H., Hoffman, E. A., Sonka, M., 2004. Automated segmentation of pulmonary vascular tree from 3D CT images. In: Proc. SPIE Med. Imaging. Vol. 5369. pp. 107–116.

Siddiqi, K., Bouix, S., Tannenbaum, A., Zucker, S., 2002. Hamilton-Jacobi skeletons. Int. J. Comput. Vision 48 (3), 215–231.

Silverman, B., 1986. Density estimation for statistics and data analysis. Chapman & Hall.

Simandl, M., Straka, O., 2002. Nonlinear estimation by particle filters and Cramer-Rao bound. Proc. IFAC.

Socher, R., Barbu, A., Comaniciu, D., 2008. A learning based hierarchical model for vessel segmentation. In: Proc. IEEE Int. Symp. Biom. Imaging. pp. 1055–1058.

Sonka, M., Fitzpatrick, J., 2000. Handbook of Medical Imaging: Medical Image Processing and Analysis. No. 2. SPIE Press.

Sonka, M., Hlavac, V., Boyle, R., Ray, L., 1996. Image Processing, Analysis and Machine Vision. Journal of Electronic Imaging 5, 423.

Soto, A., 2005. Self adaptive particle filter. In: Int. Conf. Artif. Intell. Vol. 19. p. 1398.

Stefancik, R. M., Sonka, M., 2001. Highly automated segmentation of arterial and venous trees from three-dimensional magnetic resonance angiography (MRA). Int. J. Cardiovasc. Imaging 17 (1), 48.

Steger, C., 1996. Extracting curvilinear structures: A differential geometric approach. In: Proc. Eur. Conf. Comput. Vision. Vol. 1. pp. 630–641.

Steger, C., 1998. An unbiased detector of curvilinear structures. IEEE Trans. Pattern Anal. Mach. Intell. 20 (2), 113–125.

Straka, O., Simandl, M., 2004. Sample size adaptation for particle filters. In: Proc. IFAC Symp. Automatic Control in Aerospace. pp. 444–449.

Straka, O., Simandl, M., 2006. Particle filter adaptation based on efficient sample size. In: Proc. IFAC Symp. System Identification. pp. 991–996.

Sun, Y., Lucariello, R., Chiaramida, S., 1995. Directional low-pass filtering for improved accuracy and reproducibility of stenosis quantification in coronary angiograms. In: IEEE Trans. Med. Imaging. Vol. 14. pp. 242–248.

Suri, J., Liu, K., Reden, L., Laxminarayan, S., 2002a. A review on MR vascular image processing algorithms: acquisition and prefiltering: part I. IEEE Trans. Inf. Technol. Biomed. 6 (4), 324–337.

Suri, J., Liu, K., Reden, L., Laxminarayan, S., 2002b. A review on MR vascular image processing: skeleton versus nonskeleton approaches: part II. IEEE Trans. Inf. Technol. Biomed. 6 (4), 338–350.

Suri, J., Liu, K., Reden, L., Laxminarayan, S., 2002c. White and black blood volumetric angiographic filtering: ellipsoidal scale-space approach. IEEE Trans. Inf. Technol. Biomed. 6 (2), 142–158.

Szymczak, A., 2008. Vessel tracking by connecting the dots. The Insight Journal.

Szymczak, A., Tannenbaum, A., Mischaikow, K., 2005. Coronary vessel cores from 3D imagery: a topological approach. In: Proc. SPIE Med. Imaging. Vol. 5747. pp. 505–513.

Tankyevych, O., Talbot, H., Dokladal, P., 2008. Curvilinear morpho-hessian filter. In: Proc. IEEE Int. Symp. Biom. Imaging. pp. 1011–1014.

Tek, H., Comaniciu, D., Williams, J., 2001. Vessel detection by mean-shift based ray propagation. In: Proc. IEEE Math. Methods Biomed. Image Anal. p. 228.

Tek, H., Gulsun, M., Laguitton, S., Grady, L., Lesage, D., Funka-Lea, G., 2008. Automatic coronary tree modeling. The Insight Journal.

ter Haar Romeny, B., August 2003. Front-End Vision and Multi-Scale Image Analysis: Multi-scale Computer Vision Theory and Applications, Written in Mathematica. Kluwer.

Terzopoulos, D., Witkin, A. P., Kass, M., 1988. Constraints on deformable models: Recovering 3D shape and nonrigid motion. Artif. Intell. 36 (1), 91–123.

Toledo, R., Orriols, X., Binefa, X., Radeva, P., Vitria, J., Villanueva, J., 2000a. Tracking elongated structures using statistical snakes. In: Proc. IEEE Int. Conf. Pattern Recognit. Vol. 1. pp. 157–162.

Toledo, R., Orriols, X., Radeva, P., Binefa, X., Vitrià, J., Morales, C. C., Villanueva, J. J., 2000b. Eigensnakes for vessel segmentation in angiography. In: Proc. IEEE Int. Conf. Pattern Recognit. Vol. 4. pp. 340–343.

Tschirren, J., Hoffman, E., McLennan, G., Sonka, M., 2005. Intrathoracic airway trees: segmentation and airway morphology analysis from low-dose CT scans. IEEE Trans. Med. Imaging 24 (12), 1529–1539.

Tsitsiklis, J., sep 1995. Efficient algorithms for globally optimal trajectories. In: IEEE Trans. Automation and Control. Vol. 40. pp. 1528–1538.

Tuchschmid, S., Weber, O., Martin, A., , Boesiger, P., 2005. Coroviz: Visualization of 3D whole-heart coronary artery MRA data. J. Cardiovasc. Magn. Reson. 7, 220–221.

Tyrrell, J., di Tomaso, E., Fuja, D., Tong, R., Kozak, K., Jain, R., Roysam, B., 2007. Robust 3-D Modeling of Vasculature Imagery Using Superellipsoids. IEEE Trans. Med. Imaging 26 (2), 223–237.

Udupa, J., Saha, P., Lotufo, R., 2002. Relative fuzzy connectedness and object definition: theory, algorithms, and applications in image segmentation. IEEE Trans. Pattern Anal. Mach. Intell. 24 (11), 1485–1500.

van Bemmel, C., Wink, O., Verdonck, B., Viergever, M., Niessen, W., 2003. Blood pool contrast-enhanced MRA: improved arterial visualization in the steady state. IEEE Trans. Med. Imaging 22 (5), 645–652.

Van der Merwe, R., Doucet, A., De Freitas, N., Wan, E., 2001. The unscented particle filter. Adv. Neural Inform. Process. Syst., 584–590.

Vasilevskiy, A., Siddiqi, K., 2002. Flux maximizing geometric flows. IEEE Trans. Pattern Anal. Mach. Intell. 24 (12), 1565–1578.

Vedaldi, A., Soatto, S., 2008. Quick shift and kernel methods for mode seeking. In: Proc. Eur. Conf. Comp. Vis. Springer.

Verdonck, B., Bloch, I., Maitre, H., Vandermeulen, D., Suetens, P., Marchal, G., 1995. Blood vessel segmentation and visualization in 3D MR and spiral CT angiography. Comput. Assist. Radiol. and Surg., 177–182.

Verdonck, B., Bloch, I., Maitre, H., Vandermeulen, D., Suetens, P., Marchal, G., 1996. Accurate segmentation of blood vessels from 3D medical images. In: Proc. IEEE Int. Conf. Image Process. Vol. 3. pp. 311–314.

Wan, M., Liang, Z., Ke, Q., Hong, L., Bitter, I., Kaufman, A., Co, B., Seattle, W., 2002a. Automatic centerline extraction for virtual colonoscopy. IEEE Trans. Med. Imaging 21 (12), 1450–1460.

Wan, S. Y., Kiraly, A. P., Ritman, E. L., Higgins, W. E., 2000. Extraction of the hepatic vasculature in rats using 3-D micro-CT images. In: IEEE Trans Med Imaging. 2000. Vol. 19(9). pp. 964–971.

Wan, S. Y., Ritman, E. L., Higgins, W. E., 2002b. Multi-generational analysis and visualization of the vascular tree in 3D micro-CT images. In: Computers in Biology and Medicine. Vol. 32(2). pp. 55–71.

Wand, M., Jones, M., 1995. Kernel smoothing. Chapman & Hall/CRC.

Weickert, J., 1999. Coherence-enhancing diffusion filtering. Int. J. Comput. Vision 31 (2), 111–127.

Welch, G., Bishop, G., 2001. An introduction to the Kalman filter. In: ACM Siggraph course 8. Annual Conf. Computer Graphics and Interactive Techniques.

Wenninger, M., 1999. Spherical Models. Dover.

Wesarg, S., Firle, E., 2004. Segmentation of vessels: The corkscrew algorithm. In: Proc. SPIE Med. Imaging. pp. 1609–1620.

West, M., Harrison, J., 1997. Bayesian Forecasting and Dynamic Models, 2nd Edition. Springer Series in Statistics. Springer.

Westin, C., Wigstroem, L., Loock, T., Sjoeqvist, L., Kikinis, R., Knutsson, H., 2001. Three-dimensional adaptive filtering in magnetic resonance angiography. J. Magn. Reson. Imaging 14 (1), 63–71.

Wilkinson, M. H. F., Westenberg, M. A., 2001. Shape preserving filament enhancement filtering. In: Proc. Med. Image Comput. Comput. Assist. Interv. pp. 770–777.

Williams, J., Johnstone, J., Wolff, L., 1997. Rational discrete generalized cylinders and their application to shape recovery in medical images. Proc. IEEE Conf. Comput. Vision and Pattern Recognit.

Wilson, D. L., Noble, J. A., 1999. An adaptive segmentation algorithm for time-of-flight MRA data. IEEE Trans. Med. Imaging 18 (10), 938–945.

Wink, O., Niessen, W., Viergever, M., 1998. Fast quantification of abdominal aortic aneurysms from CTA volumes. In: Proc. Med. Image Comput. Comput. Assist. Interv. pp. 138–145.

Wink, O., Niessen, W., Viergever, M., 2004. Multiscale vessel tracking. IEEE Trans. Med. Imaging 23 (1), 130–133.

Wink, O., Niessen, W. J., Verdonck, B., Viergever, M. A., 2001. Vessel axis determination using wave front propagation analysis. In: Proc. Med. Image Comput. Comput. Assist. Interv. pp. 845–853.

Wink, O., Niessen, W. J., Viergever, M. A., 2000a. Fast delineation and visualization of vessels in 3D angiographic images. IEEE Trans. Med. Imaging 19 (4), 337–346.

Wink, O., Niessen, W. J., Viergever, M. A., 2000b. Minimum cost path determination using a simple heuristic function. In: Proc. IEEE Int. Conf. Pattern Recognit. Vol. 7. pp. 010–013.

Wong, W., Chung, A., 2006. Augmented vessels for quantitative analysis of vascular abnormalities and endovascular treatment planning. IEEE Trans. Med. Imaging 2006, 665–84.

Wong, W., Chung, A., Yu, S., 2004. Local orientation smoothness prior for vascular segmentation of angiography. In: Proc. Eur. Conf. Comput. Vision. pp. 353–365.

World Health Organization, 2008. The top ten causes of death - fact sheet n310.

Wörz, S., Rohr, K., 2004. A new 3D parametric intensity model for accurate segmentation and quantification of human vessels. In: Proc. Med. Image Comput. Comput. Assist. Interv. pp. 491–499.

Wörz, S., Rohr, K., 2007. Segmentation and quantification of human vessels using a 3-d cylindrical intensity model. In: Proc. IEEE Trans. Image Process. Vol. 16. pp. 1994–2004.

Wörz, S., Rohr, K., 2008. CramÈr-rao bounds for estimating the position and width of 3d tubular structures and analysis of thin structures with application to vascular images. J. M. Imaging and Vision 30, 167–180.

Wu, C., Agam, G., Roy, A. S., III, S. G. A., 2004. Regulated morphology approach to fuzzy shape analysis with application to blood vessel extraction in thoracic CT scans. In: Proc. SPIE Med. Imaging. Vol. 5370. pp. 1262–1270.

XIP, 2009. The eXtensible Imaging Platform (XIP). http://openxip.org/.

Xu, C., Prince, J., 1998. Snakes, shapes, and gradient vector flow. IEEE Trans. Image Process. 7, 359–369.

Xu, M., Pycock, D., 1999. A scale-space medialness transform based on boundary concordance voting. J. M. Imaging and Vision 11 (3), 277–299.

Yan, J., Zhuang, T., 2003. Applying improved fast marching method to endocardial boundary detection in echocardiographic images. Pattern Recognit. Letters 24 (15), 2777–2784.

Yan, P., Kassim, A. A., 2005. MRA image segmentation with capillary active contour. In: Proc. Med. Image Comput. Comput. Assist. Interv. pp. 51–58.

Yang, C., Duraiswami, R., Gumerov, N., Davis, L., 2003. Improved fast Gauss transform and efficient kernel density estimation. In: Proc. IEEE Int. Conf. Comput. Vision. pp. 664–671.

Yang, Y., Tannenbaum, A., Giddens, D., 2004. Knowledge-based 3D segmentation and reconstruction of coronary arteries using CT images. In: Proc. IEEE Eng. Med. Biol. Soc. pp. 1664–1666.

Yatziv, L., Bartesaghi, A., Sapiro, G., 2006. O (N) implementation of the fast marching algorithm. J. Comput. Phys. 212 (2), 393–399.

Yi, J., Ra, J. B., 2003. A locally adaptive region growing algorithm for vascular segmentation. Int. J. Imaging Systems and Technol. 13 (4), 208–214.

Yim, P., Cebral, J., Mullick, R., Marcos, H., Choyke, P., 2001. Vessel surface reconstruction with a tubular deformable model. IEEE Trans. Med. Imaging 20 (12), 1411–1421.

Yim, P. J., Choyke, P. L., Summers, R. M., 2000. Gray-scale skeletonization of small vessels in magnetic resonance angiography. IEEE Trans Med Imaging 19 (6), 568–76.

Yim, P. J., Kayton, M., Miller, W., Libutti, S., Choyke, P. L., 2003. Automated detection of blood vessels using dynamic programming. Pattern Recognit. Letters 24 (14), 2471–2478.

Young, S., Pekar, V., Weese, J., 2001. Vessel segmentation for visualization of MRA with blood pool contrast agent. In: Proc. Med. Image Comput. Comput. Assist. Interv. pp. 491–498.

Yuille, A., Coughlan, J., 1997. Twenty questions, focus of attention, and A*: A theoretical comparison of optimization strategies. In: Energy Minimization Methods in Comput. Vision and Pattern Recognit. p. 197.

Yuille, A., Coughlan, J., 2000. Fundamental limits of Bayesian inference: order parameters and phase transitions for road tracking. IEEE Trans. Pattern Anal. Mach. Intell. 22 (2), 160–173.

Zahlten, C., 1995. Reconstruction of branching blood vessels from CT-data. Visualization in Scientific Computing, 41–52.

Zambal, S., Hladuvka, J., Kanitsar, A., Buhler, K., 2008. Shape and appearance models for automatic coronary artery tracking. The Insight Journal 4.

Zamir, M., Brown, N., 1982. Arterial branching in various parts of the cardiovascular system. Am. J. Anat. 163 (4), 295–307.

Zamir, M., Brown, N., 1983. Internal geometry of arterial bifurcations. J. Biomech. 16 (10), 857–63.

Zana, F., Klein, J.-C., 2001. Segmentation of vessel-like patterns using mathematical morphology and curvature evaluation. IEEE Trans. Image Process. 10 (7), 1010–1019.

Zerroug, M., Nevatia, R., 1996. Three-dimensional descriptions based on the analysis of the invariant and quasi-invariant properties of some curved-axis generalized cylinders. IEEE Trans. Pattern Anal. Mach. Intell. 18 (3), 237–253.