



HAL
open science

Image processing for semantic analysis of the coronary interventions in cardiology

Ketan Bacchuwar

► **To cite this version:**

Ketan Bacchuwar. Image processing for semantic analysis of the coronary interventions in cardiology. Document and Text Processing. Université Paris-Est, 2018. English. NNT : 2018PESC1074 . tel-01981918

HAL Id: tel-01981918

<https://pastel.hal.science/tel-01981918>

Submitted on 15 Jan 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Université Paris-Est
École doctorale MSTIC

Image processing for semantic analysis of the percutaneous coronary interventions

Traitement d'images pour l'analyse sémantique
des interventions coronariennes percutanées

par

Ketan Bacchuwar

Thèse soumise pour l'obtention du grade de docteur en informatique
de l'université Paris Est

au

laboratoire d'informatique Gaspard-Monge
ESIEE

Jury

Isabelle Bloch - Rapporteur
Grégoire Malandain - Rapporteur
Nassir Navab - Examineur
Pierre Jannin - Examineur
Maxime Sermesant - Examineur
Régis Vaillant - Examineur
Laurent Najman - Directeur de thèse
Jean Cousty - Co-directeur de thèse

Acknowledgements

There are no words to express my gratitude for my supervisors, in no particular order, Régis Vaillant, Laurent Najman and Jean Cousty, who guided me throughout my Ph.D. I appreciate a lot their enthusiasm, passion and devotion for science. Their scientific aptitude and ability to always keep encouraging and bringing the best from their student is unmatched. I will always be grateful for the efforts they put to groom me all the way till now and all the patience they kept with me.

It is a great pleasure to acknowledge my deepest gratitude to examiners Pierre Jannin, Nassir Navab and Maxime Sermesant for accepting to be a part of the jury. I am extremely fortunate to have Grégoire Malandain and Isabelle Bloch as reviewers, who invested their valuable time in critically examining my thesis and providing suggestions that contributed in significantly improving the quality of my thesis.

I deeply appreciate the stimulating discussions during the ECM team meetings with Liliane, Régis and Emmanuelle. I would like to thank my colleagues at GE, especially Marie, Romain, Maxime, Francois, Cyril, Yves and Charlotte. Special thanks to Aymeric, Thomas and Vincent for all the discussions at 'Waffle Time' breaks! It was always fun to exchange our cultural background during our discussions on endless topics. It has been a great pleasure to work with an understanding manager like Valerie, who was always encouraging and providing resources to learn.

I want to thank colleagues at ESIEE Paris, Kacper, Alberto, Bruno, Julie, Clara, Diane, Elodie, and others for their support and help. It will be unfair to not mention the constant moral support from close friends Jay, Pooja, Prakhar, Suraj, Yijun, Abhishek, Yelena, Sandeep, Manish and all the friends at India House.

Lastly but most importantly I want to acknowledge all those endless efforts and sacrifices that my parents did every single day to support me throughout my life. It's impossible to pin downs words of appreciation for the constant support from my parents and brother Sanket. I will always be indebted for making me who I am!

Résumé long

Résumé

Les interventions coronariennes percutanées (ICP) sont réalisées à l'aide de la diffusion en temps réel d'images radiographiques dans une suite interventionnelle. La modélisation de ces procédures ICP pour aider le praticien implique la compréhension des différentes phases de la procédure ICP par la machine d'intervention, qui peut être utilisée pour optimiser la dose de rayons X et l'agent de contraste. Pour atteindre cet objectif, l'une des tâches importantes consiste à segmenter différents outils d'intervention dans les flux d'images fluoroscopiques et à en déduire des informations sémantiques. L'arbre des composants, un puissant outil morphologique mathématique, constitue la base des méthodes de segmentation proposées. Nous présentons ce travail en deux parties: 1) la segmentation du cathéter vide à faible contraste, et 2) la segmentation de la pointe du guide et le suivi de la détection du vaisseau d'intervention. Nous présentons une nouvelle méthode de segmentation basée sur l'espace à plusieurs échelles pour détecter des objets faiblement contrastés comme un cathéter vide. Pour la dernière partie, nous présentons la segmentation de la pointe du guide avec le filtrage basé sur l'arbre de composants et proposons un algorithme pour suivre sémantiquement la pointe segmentée pour déterminer le vaisseau d'intervention.

Mots Clef

Radiographie, morphologie mathématique, imagerie médicale, segmentation, suivi

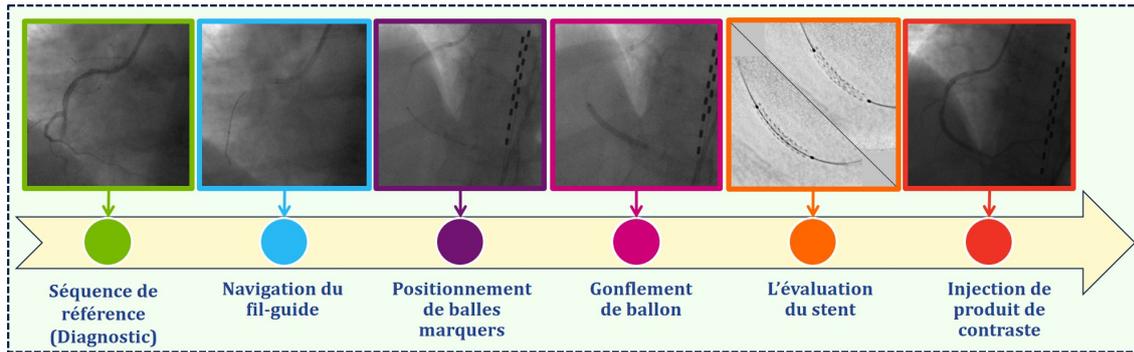


Figure 1: Phases de la procédure ICP.

A Introduction

L'intervention coronarienne percutanée (ICP) est une procédure employée pour le traitement de la sténose de l'artère coronaire. Une sténose est un rétrécissement d'une coronaire, ce qui arrive en cas de maladie coronarienne, une pathologie commune avec son état aigu la crise cardiaque pouvant conduire à la mort l'ICP est une procédure très mature reposant sur le déploiement d'un stent ayant la forme de l'artère à l'endroit de la sténose. Les outils utilisés sont navigués dans le réseau artériel du patient. Le guidage est réalisé en visualisant ces outils à travers un flux d'images radiographiques en temps réel. La contribution clinique de ces procédures est très reconnue: c'est ainsi qu'environ 175 000 stents (prothèses endovasculaires intra-coronaires) sont déployés par an en France et 2,3 millions dans les pays de l'OECD [OECD, 2011]. Par conséquent, l'optimisation du coût et de la qualité du résultat est de la plus haute importance. En fait, la quantité élevée de radiations ionisantes et d'agent de contraste (utilisé pour visualiser la vascularisation sous rayons X) est l'une des limites de ces procédures, qui a un impact sur les patients ainsi que sur le praticien. Dans ce contexte, il est souvent observé que l'équipement d'imagerie est utilisé de façon sous-optimale car il est difficile pour le praticien d'optimiser les points de fonctionnement compte tenu des multiples tâches qu'il doit effectuer. Par conséquent, le compromis qualité d'image / dose de rayonnement / agent de contraste utilisé n'est pas toujours adéquat.

Les interactions de l'équipement d'imagerie avec son utilisateur peuvent être améliorées, en particulier via des optimisations du comportement de la machine en fonction des phases de la procédure. Obtenir des informations sur les phases de la procédure directement auprès de l'opérateur humain n'est pas acceptable du point de vue du workflow. Nous visons donc à concevoir une famille d'algorithmes de traitement d'images pour identifier la présence de différents outils d'intervention dans les images et relier cette information à des connaissances de haut niveau décrivant les phases de la procédure et les attentes des utilisateurs pour chacune d'entre elles. La compréhension de ces phases dépend en partie d'une analyse sémantique des images produites tout au long de la procédure. Nous appelons cette analyse sémantique de la procédure *modélisation de procédure ICP*. Une telle analyse d'image a été étudiée dans la vision par ordinateur comme la reconnaissance

d'activité [Turaga et al., 2008] et a été traduite dans le domaine médical pour l'application chirurgicale [Lalys and Jannin, 2014]. En cas de procédures ICP, les cliniciens ont mentionné l'importance des rapports structurés et une documentation complète des indications pour les procédures ICP [Nallamotheu et al., 2014; Sanborn et al., 2014]. Les différentes phases des procédures ICP sont représentées sur la Fig. 1.7. La détection et le suivi des outils d'intervention comme le cathéter de guidage, le cathéter, la pointe du guide, le corps du guide, les billes de repérage, le ballon, le stent sont nécessaires pour construire cette modélisation. Ainsi, la segmentation de ces outils est une brique fondamentale. Il est alors envisageable de déterminer des informations sémantiques sur les régions d'intérêt, les points de contrôle et les phases à l'aide des outils segmentés. L'apparence de ces différents objets peut varier beaucoup dans les images radiographiques projetées. Les nombreux facteurs qui contribuent à cette variation sont l'apparence peut être la taille et la forme de l'objet, la radio-opacité du matériau et l'orientation de l'outil dans le plan d'imagerie.

Nous divisons ce travail en deux parties, à savoir: 1) la segmentation du cathéter vide et 2) la détection du vaisseau d'intervention en suivant l'extrémité du guide. Les articles de la version longue de chacune de ces parties [Bacchuwar et al., 2017a,b] peuvent être lus pour plus de détails. Ces tâches impliquent la segmentation du cathéter vide et l'extrémité du fil-guide dans les images fluoroscopiques respectivement. Nous choisissons un cadre unique d'arbres de composants basés sur la morphologie mathématique, un espace de forme basé sur l'arbre, pour segmenter le cathéter vide et l'extrémité du fil-guide. Sec. B fournit un bref aperçu des opérateurs morphologiques connectés mathématiques et de l'espace de forme basé sur les arbres utilisé pour la segmentation.

B Opérateurs connectés et espaces de formes basés sur des arbres

En morphologie mathématique, les opérateurs connectés [Serra and Salembier, 1993; Salembier and Serra, 1995; Salembier and Wilkinson, 2009] sont des outils de filtrage qui peuvent être définis pour agir sur des ensembles. Dans le cas des images en niveaux de gris, elles ne modifient pas les pixels individuels mais agissent sur le niveau des composants connectés, appelés zones plates [Salembier and Serra, 1995]. Les opérateurs connectés fonctionnent en fusionnant les zones plates adjacentes, ce qui garantit qu'elles ne créent ni de nouveaux contours ni ne modifient l'emplacement ou la forme des contours dans l'image d'entrée. Les opérateurs connectés sont définis à l'origine par le concept de partition de composants connectés. Soit \mathbf{P} une partition et $\mathbf{P}(p)$ une région de la partition contenant le pixel p . Nous pouvons créer une relation d'ordre partiel entre partitions: étant donné deux partitions \mathbf{P}_1 et \mathbf{P}_2 de l'ensemble V , partition \mathbf{P}_1 est dit être *un raffinement de* \mathbf{P}_2 (écrit comme $\mathbf{P}_1 \sqsubseteq \mathbf{P}_2$) si toute région de \mathbf{P}_1 est incluse dans une région de \mathbf{P}_2 (*c-à-d* $\mathbf{P}_1(x) \subseteq \mathbf{P}_2(x)$). Les opérateurs connectés sont donc définis comme,

Definition B.1. Un opérateur ψ travaillant sur une image en niveaux de gris f est

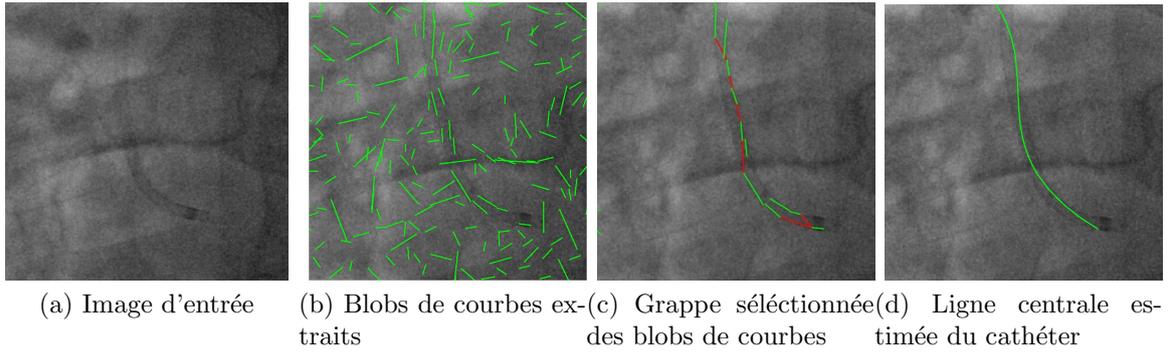


Figure 2: Segmentation d'un cathéter vide.

dit *connexe* si la partition des zones plates de l'image d'entrée f est un raffinement de la partition des zones plates de $\psi(f)$.

Une implémentation populaire des opérateurs connexes repose sur une représentation arborescente de l'image. Le filtrage consiste alors à concevoir un attribut caractérisant la forme des composantes ou le degré de ressemblance entre la forme des composantes et la forme attendue.

L'arborescence des composantes est la plus simple des représentations hiérarchiques des images. Cette représentation est composée des régions de l'image (*c.à.d.* des composantes connexes) de la plus petite à la plus grande, deux régions différentes étant soit incluses l'une dans l'autre, soit disjointes. Ces régions peuvent donc être organisées sous la forme d'une structure arborescente, chaque région R formant un nœud \mathcal{N} de l'arbre. Le nœud racine désigne le domaine entier de l'image. Deux régions distinctes R_1 et R_2 , représentées par deux nœuds \mathcal{N}_1 et \mathcal{N}_2 , ont une propriété d'inclusion: $R_1 \cap R_2 \neq \emptyset \rightarrow R_1 \subset R_2$ or $R_2 \subset R_1$. Si $R_1 \subset R_2$, et qu'il n'existe pas de région R_3 telle que $R_1 \subset R_3 \subset R_2$, on dit que \mathcal{N}_2 est parent de \mathcal{N}_1 . Hormis pour le nœud racine, chaque nœud \mathcal{N} de l'arbre a un nœud parent unique \mathcal{N}_p , et il existe une arête $(\mathcal{N}, \mathcal{N}_p)$ les reliant, reflétant cette relation de parenté.

Cette notion nous aide à définir des espaces de formes arborescentes $S_{\mathcal{T}}$ [Xu et al., 2016] pour toute représentation d'image basée sur l'arbre:

Definition B.2. Espace arborescent de formes. Un espace de forme d'arbre $S_{\mathcal{T}}$ est défini comme un ensemble R_i de régions, qui peut être organisé en une structure arborescente \mathcal{T} pour laquelle le nœud racine représente R_0 , le domaine entier de l'image et deux régions R_1 et R_2 sont soit incluses, soit disjointes. Le voisinage de l'espace de forme est défini par la relation d'inclusion entre les deux régions.

Dans ce travail, nous nous intéressons à des espaces de formes, qui munis d'une certaine fonction de pondération (*c.à.d.*, un poids est associé à chacune des formes) forment une représentation équivalente de l'image dans le sens où l'image peut être reconstruite à partir de l'espace de forme et de sa pondération. Nous décrivons brièvement ci-dessous deux types de représentations arborescentes d'images utilisées dans ce travail: **Min tree**. Le *min tree* [Salembier and Wilkinson, 2009] structure les composants connectés des ensembles de niveau inférieur de l'image en niveaux de

gris en fonction de la relation d'inclusion. Une image en niveaux de gris f , lorsqu'elle est seuillée dans un ordre croissant à chaque niveau de gris possible allant de h_{min} à h_{max} , donne une pile d'ensembles de niveaux inclus (inférieurs). Chaque ensemble de niveaux peut être partitionné en composants connectés lorsque le domaine d'image est structuré comme un graphe d'adjacence de pixels (nous considérons la relation 4-adjacence). Dans ce paramètre, deux composants connectés à deux seuils successifs sont inclus ou disjoints. L'ensemble de tous les composants connectés est un arbre dirigé appelé *min tree* $\mathcal{T}^>$ de l'image f .

Hiérarchie par zones quasi-plates. L'espace de forme appelé *hiérarchie des zones quasi-plates* [Meyer and Maragos, 2000] est la hiérarchie induite par les ensembles de niveaux d'un graphe aux arêtes pondérées. Par exemple, on peut pondérer les arêtes du graphe 4-connexe de l'image par la différence des niveaux de gris entre les deux pixels de chaque arête. Une différence importante avec le min tree est que chaque ensemble de niveaux d'une hiérarchie de zones quasi-plates forme une partition du domaine de l'image.

Segmentation de cathéter vide Un cathéter de guidage est un outil présent tout au long de la procédure ICP. Il contribue à l'information sémantique de façon significative, puisqu'il est le premier outil présent dans le champ de vision. Il est positionné à l'ostia de l'artère coronaire gauche ou droite pour le reste de la procédure. C'est le conduit pour tous les autres outils / dispositifs qui seront amenés dans le vaisseau où la lésion est présente. L'agent de contraste est également injecté dans le système vasculaire à travers le cathéter de guidage. Ainsi, sa segmentation peut aider à la modélisation de la procédure ICP pour déterminer les événements / phases de l'arrivée et du retrait des autres dispositifs (guide, les marqueurs des ballons d'angioplastie).

Plusieurs méthodes ont été proposées dans la littérature pour détecter différents types de cathéters comme les cathéters d'électrophysiologie (EP) [Baur et al., 2016], les cathéters de sinus coronaires [Wu et al., 2011], les cathéters en queue de cochon [Lessard et al., 2015]. Les inconvénients majeurs de ces méthodes sont: i) la première image doit être annotée manuellement et ii) la longueur et la courbure du guide et du cathéter ne doivent pas beaucoup changer pendant la séquence. Récemment, [Ambrosini et al., 2017] a proposé une segmentation automatique à base de réseau neuronal convolutif profond basée sur un guide et un cathéter. Toutes ces approches se concentrent sur le suivi du fil-guide avec le corps du cathéter. Dans ce travail, nous abordons la tâche de la segmentation du cathéter de guidage vide dans les images fluoroscopiques. En raison du faible contraste dans les images fluoroscopiques, le cathéter vide apparaît comme une structure faiblement contrastée avec deux bords parallèles et partiellement déconnectés car il s'agit juste d'un tube vide fait d'un matériau avec peu de radio-opacité.

B.1 Méthode

Dans ce travail, nous utilisons d'abord l'espace d'échelle de niveau pour identifier les blobs de courbe (Sec. 3.4.1), qui sont de petites régions sombres persistantes qui font potentiellement partie du cathéter vide. Nous proposons un nouvel espace

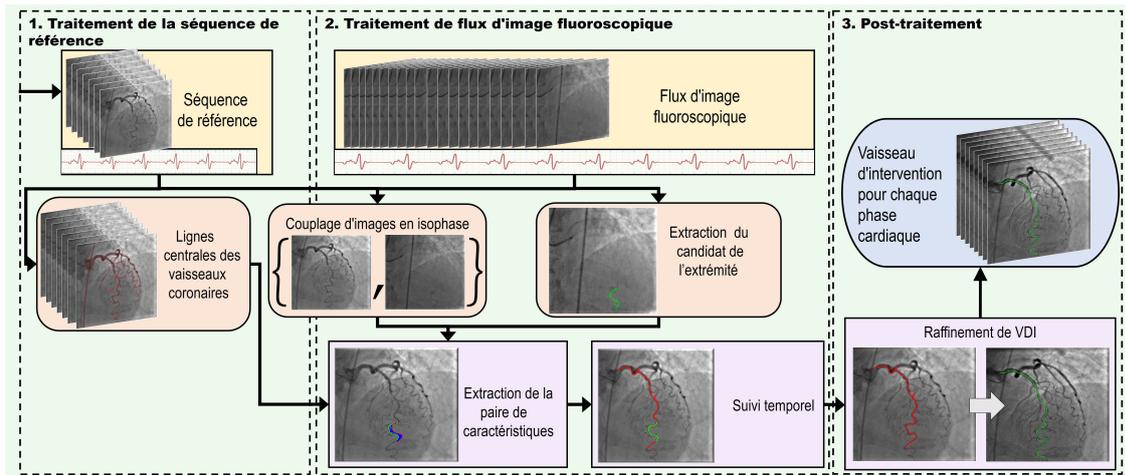


Figure 3: Flux de travail du système de détection et suivi du vaisseau d'intervention.

d'échelle basé sur un graphe structurel, sous la forme d'une hiérarchie, *c.à.d* un arbre construit sur les blobs de la courbe (Sec. 3.4.2). Nous analysons ce second arbre avec les mêmes techniques que le premier, et nous conservons les structures les plus persistantes dans ce second espace d'échelle comme segmentation finale. Si le premier arbre présente les structures profondes des points critiques, le deuxième arbre met en avant les structures d'intérêt encore plus profondes, que nous appelons *composants connectés profonds*.

Extraction de blob de courbe Nous construisons d'abord le min tree de l'image d'entrée. Le min tree considère seulement les régions sombres connectées de l'image et les blobs de la courbe apparaissent comme des régions à différents niveaux de cet arbre (*c.à.d*, à différentes échelles.) Ensuite, l'ensemble des blobs de la courbe est inclus dans une coupe horizontale de ce min tree. On attribue à chaque composante du min tree des attributs caractérisant sa forme et ses propriétés structurelles. Nous utilisons les attributs d'aire et d'élongation pour filtrer les blobs de courbe, en concevant un critère de sélection basé sur ces attributs. Les régions retenues forment l'ensemble des blobs de la courbe. En fonction de la structure du min tree, un filtrage est effectué pour conserver les éléments ayant la plus grande surface.

Groupe ment de blob de courbe Cette section présente l'idée principale de la détection de cathéter vide, *c.à.d* le regroupement de blobs de courbe dans un espace d'échelle structurel. Les blobs de courbe extraits à l'étape précédente peuvent ou non être des régions des bords du cathéter. En analysant individuellement un blob donné, il est difficile de décider s'il fait partie d'un cathéter, parce que l'information contextuelle est manquante. Nous les considérons donc dans un espace commun et définissons un poids pour chaque paire de blobs de courbe, appelé *poids de la paire blob*, formant ainsi un graphe pondéré par les bords. Nous proposons de le construire en combinant différents poids élémentaires, a) un poids Euclidien et un poids d'alignement basé sur des informations géométriques; et b) le poids du profil basé sur l'intensité de l'image. Ainsi, les blobs connectés avec un poids de paire blob forment un graphe pondéré par les bords, dont les ensembles de niveaux forment une

hiérarchie de zones quasi-plates. Une *composante connexe profonde* est un élément de cette hiérarchie, qui est un cluster de blobs de courbe. Nous explorons donc cette hiérarchie de la même manière que le min tree en concevant des attributs qui caractérisent les propriétés du cathéter vide. La segmentation résultante est formée par la ligne médiane estimée du cathéter, qui est une courbe polynomiale ajustée à la grappe de blobs maximisant le score de probabilité d'être cathéter.

B.2 Évaluation et résultats

Nous évaluons notre algorithme de segmentation de cathéter vide en utilisant un ensemble de données d'images fluoroscopiques d'environ 1250, appartenant à des 10 séquences prélevées sur des examens de 6 patients. Les images considérées représentent une grande variabilité en raison de l'indice de masse corporelle (IMC) des patients, des niveaux de bruit, de différents antécédents anatomiques, de la présence occasionnelle de sondes de stimulation, de stents, d'agrafes et de fils sternaux. À la suite de l'évaluation d'une base de données de ces images fluoroscopiques, nous obtenons de très bonnes performances de segmentation qualitatives et quantitatives, avec une précision moyenne et des rappels de 83,85% et 67,87% respectivement.

C Vaisseau d'intervention détection dynamique

Le *Vaisseau D'Intervention* (VDI) est une branche de l'arbre des vaisseaux coronaires entre l'ostia et l'extrémité distale du vaisseau à travers la lésion coronaire. Dans cette partie, nous présentons un algorithme de suivi entièrement automatique, appelé Vaisseau D'Intervention Détection Dynamique (VDIDD), pour détecter le vaisseau d'intervention, traité pendant la procédure, en combinant les informations des images de vaisseau avec injection d'agent de contraste et des images fluoroscopiques acquises navigation où le VDI n'est pas visible. Un travail important existe dans la littérature avec un but différent et un flux de travail que notre but. Ce travail a été fait pour co-localiser entre différentes modalités, comme les séquences angiographiques radiographiques avec échographie intra-vasculaire (IVUS) [Frimerman et al., 2016] ou avec la tomographie par cohérence optique (OCT) [Koyama et al., 2015] qui sont impliqués dans certaines des procédures ICP.

Comme mentionné dans la section B, le cathéter est placé à l'ostia de l'arbre coronaire gauche et droit, préparant ainsi la première phase de la procédure ICP, phase de diagnostic du vaisseau (voir Fig. 1.7). Dans une étape suivante, les lésions coronaires sont traitées en naviguant un fil guide dans le vaisseau d'intervention, suivi par l'implantation d'une endoprothèse au niveau de la lésion dans le VDI. Nous combinons l'information à ces phases initiales à partir de deux séquences d'images radiographiques différentes: i) les images de ciné de *séquence de référence*, un sous-ensemble de séquence angiographique, qui sont injectées avec un agent de contraste pour décrire le système vasculaire complet; ii) les images fluoroscopiques de *flux d'image fluoroscopique*, qui sont acquises lors de la navigation du fil-guide. Notre cadre de suivi vise à déterminer la position correspondante de la pointe du fil-guide dans les vaisseaux injectées. Le point crucial de l'algorithme réside dans la

création et le suivi de paires de caractéristiques cohérentes dans le temps et dans l'espace. Afin de former de telles paires, nous présentons également (i) un algorithme d'extraction d'arbre de vaisseau, (ii) une méthode robuste de segmentation de bout de fil guide, et (iii) une procédure d'appariement permettant de faire des correspondances entre pointes de fil guide et vascularisation détectées. Les différentes étapes de traitement de la méthode sont montrées dans la Fig. 4.2. Plus précisément, cet algorithme est capable de reconnaître à partir du flux d'images fluoroscopiques, l'instant correspondant à la navigation du fil-guide et de l'exploiter pour déterminer le navire d'intervention sans ajouter de contraintes au déroulement de la procédure. Afin d'atteindre cet objectif, un cadre de suivi général est expliqué dans Sec. 4.3.1 (algorithme détaillé dans [Bacchuwar et al., 2017b]).

Nous commençons par le pré-traitement et la synchronisation des données d'image par rapport aux données non-image. Le mouvement cardiaque provoque des changements de position et de forme périodiques des vaisseaux coronaires et ces déformations s'appliquent au guide qui est navigué dans la lumière de ces vaisseaux. Il vient donc que, l'apparence d'un guide dans l'image 2D projetée par rayons X change continuellement. Par conséquent, nous avons choisi d'utiliser des données ECG pour associer les images à partir de deux séquences à une même phase cardiaque, où le système vasculaire décrit et des fils de guidage ont des déformations correspondantes induite par le mouvement cardiaque. Malgré l'appariement d'images en iso-phase, il peut subsister des incohérences spatiales entre les deux images appariées en raison du mouvement de respiration, des déformations des vaisseaux dues au corps rigide du guide, des translation de table, des rotations de l'arceau qui porte la chaîne image. Nous avons également remarqué, dans certains cas, un décalage de phase ECG induite par un retard électronique de la machine ECG. Par conséquent, il est nécessaire que l'algorithme d'appariement et de suivi suive de manière cohérente le guide dans l'image de référence en prenant en compte tous ces facteurs.

C.1 Cadre de suivi général

Nous visons à obtenir le VDI en faisant une correspondance intelligente entre la séquence de référence et le flux d'image fluoroscopique. Par conséquent, nous proposons un algorithme, appelé VDIDD, qui est capable de détecter simultanément l'extrémité du fil-guide dans les images fluoroscopiques et le VDI, c'est-à-dire la branche des vaisseaux coronaires dans laquelle le fil-guide navigue actuellement. Dans une perspective plus large, l'algorithme consiste à: i) détecter des paires de caractéristiques à partir de paires d'images isophases; ii) regrouper ces paires de caractéristiques en pistes, une piste étant une séquence de paires de caractéristiques spatialement cohérentes dans le temps; iii) sélectionner la voie la plus pertinente en tant que vaisseau d'intervention détecté. Une paire de caractéristiques est composée de deux courbes correspondantes et est extraite d'une paire d'images isophases. Le premier, appelé *candidat de pointe*, correspond peut-être à l'extrémité du guide dans l'image fluoroscopique. Le second, appelé un chemin de vaisseau, est un chemin dans l'arbre des vaisseaux coronariens qui s'adapte de manière optimale au candidat de pointe associé.

L'algorithme VDIDD gère un dictionnaire de pistes, où chaque piste est une séquence de paires d'entités, avec au plus une paire d'entités par image du flux d'image fluoroscopique. Pour chaque paire d'images en isophase, l'algorithme attribue de façon optimale chaque paire de caractéristiques détectées à l'une des pistes existantes. Un coût appelé *coût de assignation de piste* (décrit dans la section 4.3.3) est considéré comme assignant de manière optimale une paire de caractéristiques à une piste. S'il n'existe pas de piste adéquate pour une paire de caractéristiques, une nouvelle piste est initialisée. Une fois que toutes les images du flux d'images fluoroscopiques sont traitées, la piste la plus longue (, *c.à.d* la piste avec le nombre maximum de paires de caractéristiques) est sélectionnée en tant que VDI.

C.2 Extraction de paires de caractéristiques

Comme mentionné précédemment, l'extraction de paires de caractéristiques consiste à suivre les trois étapes suivantes:

Extraction de ligne centrale du vaisseau. Chaque image de la séquence de référence est pré-traitée avec fermeture morphologique pour estimer le fond anatomique. Pour améliorer les vaisseaux dans l'image soustraite résultante, une technique basée sur Hessian [Krissian et al., 2000] est adaptée qui suppose que les vaisseaux sont des structures tubulaires localement linéaires. Afin d'avoir une représentation graphique de la structure du vaisseau, nous proposons un schéma d'extraction de la ligne centrale pour obtenir les axes des vaisseaux.

Extraction du candidat de point. Afin de détecter les emplacements possibles des pointes de guide dans une image fluoroscopique, nous utilisons la même représentation de l'image, le min tree, qui était précédemment utilisé pour la segmentation du cathéter vide. L'extrémité du fil de guidage est caractérisée par des attributs de zone et d'allongement pour le filtrage dans un min tree. On observe qu'un simple seuillage sur l'attribut d'allongement n'est pas suffisant. Par conséquent, nous adoptons le cadre de mise en forme [Xu et al., 2016] qui nous permet d'extraire efficacement des composants connectés significatifs. Nous simplifions la pointe segmentée (objet de type fil) en utilisant la squelettisation [Couprie and Bertrand, 2012] pour obtenir sa ligne centrale parce que nous la considérons comme un objet 1-D. Cette simplification de l'objet 1D aide à l'étape correspondante de l'extraction de la paire de caractéristiques.

Correspondance Une étape importante dans la tâche de détection de VDI est de désigner des associations possibles de l'extrémité du fil de guidage à l'intérieur du vaisseau injecté. La tâche à accomplir dans cette étape consiste à faire correspondre et à aligner un candidat de pointe sur le graphique du vaisseau pour trouver le trajet du vaisseau correspondant. Une distance courbe-courbe doit être définie pour faire correspondre et aligner deux courbes. Nous utilisons une version discrète de la distance de Fréchet [Eiter and Mannila, 1994] car elle prend en compte l'emplacement et l'ordre des points le long des courbes en établissant une correspondance monotone d'une courbe à l'autre. Dans ce cas, nous adaptons l'appariement de la courbe de [Benseghir et al., 2015] pour effectuer la tâche d'appariement car il prend en compte la distance du Fréchet et la similarité de forme entre deux courbes.

L'ensemble obtenu de paires de caractéristiques pour chaque candidat de pointe est ensuite trié en fonction de la propriété de ressemblance de la courbe pour préférer les trajets de vaisseau avec une ressemblance de forme plus élevée avec le candidat de pointe.

C.3 Coût d'assignation de piste

Le coût d'assignation de piste est calculé comme le coût d'attribution d'une paire d'entités proposée à une piste. C'est une combinaison de trois valeurs de distance: une distance d'extrémité de pointe, une distance d'extrémité de trajet de vaisseau et une distance de graphique. Les deux premières valeurs sont définies avec la distance de l'extrémité de la courbe qui est la distance Euclidienne entre les extrémités de deux courbes (non orientées). Ensuite, la troisième valeur de distance est obtenue avec une distance de graphe qui est la longueur moyenne des trajets géodésiques (dans le graphe de ligne centrale du vaisseau) entre les points d'extrémité de deux trajets de vaisseaux donnés dans le graphique du vaisseau. Pour limiter la fragmentation de la piste, nous privilégions la longueur de piste en multipliant le coût d'assignation de piste par un facteur de longueur de piste. Les deux valeurs de distance d'extrémité de courbe représentent le décalage géométrique entre les deux candidats de pointe et les deux trajets de vaisseau respectivement. La distance de l'extrémité de la courbe de chemin et la distance du graphique conservent la cohérence dans les pistes, tandis que la distance de l'extrémité de la courbe de pointe évite que des extrémités de guide soient faussement détectées dans la piste la plus longue.

C.4 Raffinement du vaisseau d'intervention

La piste la plus longue contient le VDI détecté pour différentes phases cardiaques. A une phase cardiaque donnée, la réunion des chemins de vaisseaux associés dans cette phase constitue constituent une première proposition du VDI. Ce premier VDI proposé a souvent de petites branches, des boucles, des trous, des connexions manquantes / vaisseau coupé résultant des appariements individuels des candidats de pointe. Nous effectuons diverses opérations morphologiques pour post-traiter l'union de ces trajets de vaisseaux afin d'affiner le VDI détecté automatiquement à une phase donnée. Nous appliquons une technique de reconnexion et de fermeture de trou basée sur la transformation homotopique guidée. Le résultat final après raffinement ne contient aucune petite branche, boucle, épaisseur ou déconnexion.

C.5 Évaluation et résultats

L'ensemble de données d'évaluation comprend des séquences recueillies dans 15 cas cliniques différents, y compris 9989 images provenant de flux d'images fluoroscopiques et 140 images provenant de séquences de référence provenant de 14 patients. Ce jeu de données représente une variété de cas cliniques, y compris différentes anatomies et localisations vasculaires des lésions, angulations de l'arceau du système interventionnel, indices de masse corporelle (IMC), qualités d'image des

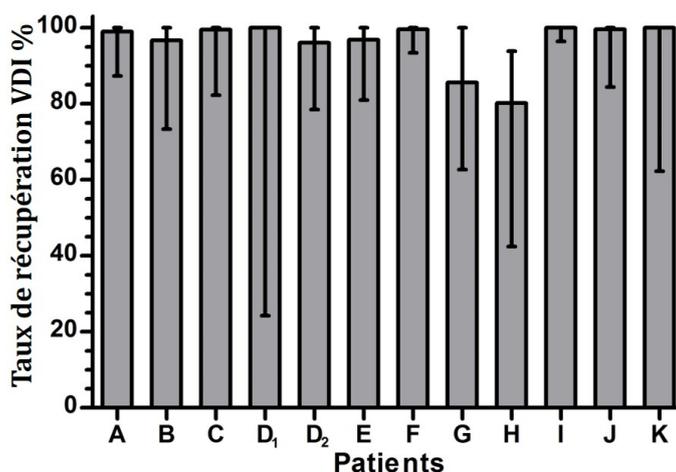


Figure 4: Les taux de récupération VDI

images fluoroscopiques et complexités de la navigation du guide. Nous avons également effectué une analyse détaillée pour analyser les cas en fonction de la qualité de l'image et des différences d'angulation de l'arceau.

Afin d'évaluer l'efficacité de l'algorithme pour identifier l'arrivée et la navigation d'un fil guide dans un flux d'image fluoroscopique, nous analysons la robustesse de l'algorithme VDIDD pour détecter les VDI dans les images sans guide dans le champ de vision. Il s'agit d'une phase de procédure pendant laquelle le système d'imagerie est employée avant que le guide n'arrive dans le champ de vue. Selon cette mesure, l'algorithme VDIDD détecte de façon robuste la pointe (d'où son arrivée) avec un taux de détection moyen pondéré de $99,05 \pm 1,61\%$.

Nous analysons des vaisseaux sur les plus longues pistes avec une mesure d'évaluation, qui estime le taux de détection en calculant la distance entre le vaisseau et la vérité terrian. Sur la base de cette mesure, les vaisseaux sont correctement détectés avec une précision moyenne de $84,09\%$. Nous avons développé également des mesures d'évaluation VDI pour analyser le VDI après raffinement. Les données ont été divisées en deux séries de 12 cas réguliers et 3 cas extrêmes basés sur l'angulation de l'arceau. Les données de 12 cas cliniques (11 patients) sont ensuite évaluées selon la mesure d'évaluation VDI comme représenté sur la Fig. 4.12, où les barres se réfèrent aux valeurs médianes, minimum et maximum du taux de récupération VDI. Pour chaque cas, ces statistiques sont calculées sur les taux de récupération VDI pour différentes phases cardiaques ($=10$). Sur notre ensemble de données, le taux moyen de récupération de VDI est $93,22\%$. L'une des observations les plus importantes est que le vaisseau d'intervention récupéré contient l'emplacement de la lésion coronaire dans 100% des cas.

D Conclusion

Dans ce travail, nous introduisons le concept de modélisation de procédure ICP comme une analyse sémantique de la progression de la procédure. Nous montrons que la segmentation des outils d'intervention est l'une des tâches les plus importantes pour une telle analyse sémantique. Nous avons démontré que les arbres de composants basés sur la morphologie mathématique sont des approches robustes pour la segmentation de plusieurs outils. La segmentation du cathéter vide basée sur l'espace à plusieurs échelles est une nouvelle méthode pour la segmentation d'objets à faible contraste. L'algorithme VDIDD pour déterminer automatiquement le vaisseau d'intervention est une nouvelle approche algorithmique dans le domaine ICP. Des résultats encourageants ont été obtenus avec un taux moyen de récupération des VDI d'environ 93,22% et une précision moyenne de détection des pointes de 99,05%, ce qui prouve la robustesse par rapport aux différentes conditions d'imagerie du patient. Nous croyons que cet algorithme de suivi peut être étendu à des applications comme le suivi des sondes IVUS. Cet algorithme a le potentiel de faire partie du logiciel embarqué par les systèmes d'imagerie à rayons X et capable de surveiller automatiquement les étapes successives de la procédure en vue d'adapter en continu le comportement du système aux besoins de l'utilisateur. Le même filtrage basé sur l'arbre min peut être utilisé pour effectuer la segmentation d'autres outils d'intervention. Ce formalisme ouvre des portes pour le développement d'applications variées comme la documentation automatique des procédures, la prédiction du début de la phase de navigation, l'amélioration des algorithmes d'amélioration des stents entre autres. Les travaux futurs comprennent l'étude de la détection des autres outils majeurs (marqueurs, ballons) en prenant appui sur le vaisseau d'intervention détecté et leur intégration dans un modèle sémantique de la procédure.

Abbreviations

BMI	B ody M ass I ndex
CABG	C oronary A rtery B ypass G raft
CNR	C ontrast-to- N oise R atio
ECG	E lectrocardiography
EP	E lectrophysiology
fps	frames p er s econd
FOV	F ield O f V iew
GT	G round T ruth
ICP	I terative C losest P oint
ICC	I terative C losest C urve
IGS	I nterventional G uiding S ystem
IQ	I mage Q uality
IVUS	I ntra V ascular U ltra S ound
LAD	L eft A nterior D escending
LCX	L eft C ircumflex A rtery
OCT	O ptical C oherence T omography
PCI	P ercutaneous C oronary I ntervention
PET	P ositron E mission T omography

QFZ	Q uasi- F lat Z one
RCA	R ight C oronary A rtery
RPD	R ight P osterior D escending A rtery
SNR	S ignal-to- N oise R atio
ToS	T ree of S hapes
TRE	T arget R egistration E rror
VOI	V essel O f I ntervention
VOIDD	V essel O f I ntervention D ynamic D etection
WHO	W orld H ealth O rganization

Contents

Résumé long	v
Abbreviations	xvii
List of Figures	xxiii
List of Tables	xxvii
1 Introduction, clinical background and motivation	1
1.1 Clinical context	2
1.1.1 Cardiovascular pathologies	2
1.1.2 Treatments	4
1.2 Detailed view on PCI	6
1.2.1 The Cathlab	6
1.2.1.1 Interventional system	6
1.2.1.2 ECG acquisition	8
1.2.2 Breakdown of a PCI procedure	8
1.2.2.1 Interventional tools	8
1.2.2.2 Phases of the procedure	9
1.3 Motivation: PCI procedure modeling and its applications	11
1.4 A word about the dataset	12
1.5 Context of the thesis	13
2 Related works	15
2.1 Segmentation and detection of interventional tools in X-ray images	16
2.2 Graph based connected filtering	18
2.2.1 General overview	18
2.2.2 Graphs and connectivity for image analysis	19
2.2.3 Connected operators	23

2.2.3.1	Mathematical morphology operators on binary and grayscale images	23
2.2.3.2	Definition of connected operators	24
2.2.3.3	Example of connected operators: attribute filters	25
2.2.4	Connectivity in weighted graphs	28
2.2.4.1	Weighted graphs	28
2.2.4.2	Stacks of graphs	29
2.2.4.3	Stacks induced by a vertex-weighted graph	30
2.2.4.4	Stacks induced by an edge-weighted graph	31
2.2.5	Connected partitions and hierarchies	32
2.2.6	Component trees: tree based shape space	36
2.2.6.1	Min/Max tree.	37
2.2.6.2	Quasi-flat zones tree	40
2.2.6.3	Algorithms for component tree construction	42
2.2.7	Filtering on component trees	43
2.2.7.1	Increasing attributes based filtering	44
2.2.7.2	Non-increasing attributes based filtering	44
2.2.7.3	Filtering with shaping	45
2.2.8	Challenges with connected filters: second order connectivity	47
2.3	Registration	48
3	Empty catheter segmentation	51
3.1	Abstract	51
3.2	Introduction	52
3.2.1	Guiding catheter vs other catheters	52
3.2.2	Related works	53
3.2.3	Guiding catheter in fluoroscopy	54
3.2.4	Proposed method at a glance	56
3.3	Scale spaces and deep connected components	56
3.4	Method	58
3.4.1	Curve blobs extraction	58
3.4.2	Curve blob clustering	60
3.4.3	Deep connected component selection	64
3.5	Segmentation Quality Evaluation	65
3.5.1	Dataset	65
3.5.2	Ground truth construction	66
3.5.3	Segmentation	66
3.5.4	Evaluation measures	66
3.6	Results and discussion	69
3.7	Conclusion of catheter segmentation	73
4	VOIDD: vessel of intervention dynamic detection and tracking	75
4.1	Introduction	76
4.1.1	Clinical interest	76
4.1.2	Related work	77
4.1.3	Major contributions	80

4.2	Data and preprocessing	81
4.2.1	Reference sequence	81
4.2.2	Fluoroscopic image stream	82
4.2.3	Iso-phase image pairing	82
4.2.4	Table panning and gantry rotation compensation	84
4.3	VOIDD algorithm	85
4.3.1	General tracking framework	87
4.3.2	Feature pairs extraction	87
4.3.2.1	Vessel centerline extraction	87
4.3.2.2	Tip candidate extraction	89
4.3.2.3	Matching	92
4.3.3	Track assignment cost	93
4.3.4	Vessel of intervention refinement	96
4.4	Dataset and evaluation measures	99
4.4.1	Dataset and expert annotations	99
4.4.2	Tip detection evaluation measure	101
4.4.3	Track evaluation measure	101
4.4.4	VOI evaluation measure	101
4.5	Results	102
4.6	Conclusion and future work	109
5	Conclusion and perspectives	111
5.1	Conclusion	111
5.2	Perspectives	113
5.2.1	Vessel of intervention for semantic applications	114
5.2.2	Cathlab of the future and Robotics in PCI	115
A	Publication list	117
A.1	Journal	117
A.2	Conferences	117
A.3	Patents	117
	Bibliography	119
	Index	135

List of Figures

1.1	Artherosclerosis explained by [National Heart, Lung and Blood Institute, 2015d]. (top) Normal artery with normal blood flow; (bottom) an artery with plaque buildup. The inset images shows cross-sections of corresponding arteries.	3
1.2	Visualization of a diseased artery under X-ray fluoroscopy (angiography) with an IGS 520 X-ray system (GE Healthcare) . The position of the stenosis is shown with an arrow.	3
1.3	An illustration of CABG [National Heart, Lung and Blood Institute, 2015b]	5
1.4	An illustration of balloon angioplasty and stenting [National Heart, Lung and Blood Institute, 2015c]	5
1.5	Innova Interventional Guiding System, IGS-520 from GE Healthcare in cathlab.	7
1.6	Comparison between fluoro and record images	7
1.7	Phases of the PCI procedure	9
2.1	An example of binary connected operator.	17
2.2	A graph	20
2.3	Adjacency relations for image analysis: (a) a two-dimensional image; (b) graph induced by 4-adjacency associated with the image (a); (c) graph induced by 8-adjacency associated with the image (a)	21
2.4	A connected operator example on grayscale image for low-level filtering task compared with classical filters: (b) low-pass (mean) filter, (c) median filter, (d) Gaussian filter, (e) opening with structuring element. (f) The grain filter does not create or move contours.	26
2.5	Attribute filters	28
2.6	An illustration of stack of graph	30
2.7	Stack induced by vertex-weighted graph	31

2.8	Stack induced by edge-weighted graph	33
2.9	An illustration of partition, partial partition and hierarchy on a set V of 9 dots. In first row, for every partition \mathbf{P}_i , each region is represented by a gray level: two dots with same gray level belong to same region. \mathbf{P}_* is the partial partition of the set V . The right subfigure in second row represents the hierarchy $\mathcal{H} = (\mathbf{P}_0, \mathbf{P}_1, \mathbf{P}_2, \mathbf{P}_3)$ as a tree structure, called as a dendogram, where inclusion relation between regions of successive partitions is represented by line segments.	36
2.10	An illustration of component tree on an image	38
2.11	An illustration of min tree as a hierarchy of partial partition	39
2.12	Illustration of quasi-flat zones tree	41
2.13	Shape-based morphological filtering scheme	46
3.1	Different types of catheters in cardiac fluoroscopy: (a) Electrophysiology (EP) catheters, abbreviations: ABL catheter, ablation catheter; CS catheter, coronary sinus catheter; RV catheter, right ventricle catheter; His catheter, (b) two types of PCI catheters including pig-tail catheter and guiding catheter. Pacing lead, a tubular tool, also looks like a catheter.	53
3.2	Guiding catheter in fluoroscopy	55
3.3	Curve blobs extraction with min tree filtering, from left to right: Input toy image; lower level-sets of input image; corresponding min tree where each node is one of the connected components of the level-sets; extracted curve blobs overlaid with green centerlines.	57
3.4	Distribution of connected components belonging to two classes (catheter and non-catheter) in the attribute space of area and elongation.	58
3.5	Curve blob extraction, (a) input fluoroscopic image and (b) green centerlines of extracted curve blobs.	59
3.6	Curve blobs clustering; from left to right: extracted curve blobs (see Figure 3.3); connection of curve blobs on structural scale-space; hierarchy of deep connected components (clusters of curve blobs at different scales); selected deep connected component overlaid on input image.	60
3.7	Blob pair weight computation, (a) w_S is the spatial weight between two blobs with blob axes Ca_1 and Ca_2 (b) vectors to compute alignment weight and (c) shows profiles \mathcal{P}_1 and \mathcal{P}_2 to compute profile weight.	62
3.8	Curve blob clustering	63
3.9	Evaluation principle to compute precision and recall. (a) Matched detection (extracted centerline); (b) Matched reference (ground truth)	67
3.10	Results of empty catheter segmentation in X-ray fluoroscopic images	68
3.11	Box plots of precision and recall at different thresholds in fluoroscopic images	70
3.12	Results of empty catheter segmentation in record and fluoroscopic images	72

3.13	Failed segmentation cases. In (a) and (b), left image shows the input image and right image shows the ground truth in green and the falsely detected catheter in red. See Section 3.6 for details.	73
4.1	Vessel of intervention definition: a branch of coronary vessel tree in which the guidewire is navigated and is between ostia (yellow region) and the distal end of the vessel (green region) across the coronary lesion (red arrow). (a) shows the VOI in injected vessel image and (a) shows the guidewire tip at the distal end of vessel.	77
4.2	Vessel of intervention detection system workflow	79
4.3	VOIDD: (from left to right) Input image f ; centerline of segmented guidewire tip; tip candidate (red) matched to vessel centerline (green) marked by pairings (blue); corresponding location (green) of guidewire tip (red) inside vessel.	82
4.4	Projective geometry in case of gantry rotation: X_1 and X_2 are projections of a 3D point X in image (detector) planes π_1 and π_2 respectively. S_1 and S_2 are the positions of the X-ray source in these two cases.	84
4.5	Background subtraction for contrasted vessels: (a) Native reference image, (b) estimated background with the help of morphological closing, (c) resulting background subtracted image.	88
4.6	Vessel centerline extraction and graph representation: (a) Native reference image with injected vessels, (b) vesselness image, (c) result of Non-Maximum Suppression (NMS) applied to vesselness image, (d) extracted vessel centerlines superimposed on the native reference image and the region marked in blue is then zoomed in (e) where the vertices of different degrees of the vessel graph \mathcal{X} . Vertices of degree 1 in yellow, vertices of degree 2 in blue, vertices of degree 3 or more in red.	88
4.7	Tip candidate extraction: (a) Native fluoroscopic image, (b) background subtracted image, (c) result of guidewire tip segmentation, (d) segmented guidewire tip after skeletonization	91
4.8	Feature pair extraction: (a) Tip candidate (red) in the iso-phase reference image with extracted centerlines (green), (b), (c), (d) and (e) corresponding feature pairs with tip candidate (in red) and vessel paths (in green)	94
4.9	Illustration of track assignment cost	95
4.10	Vessel of intervention refinement, (a,b) VOI from longest track contains small branches (shaded red), loops (shaded blue) and hole/cut vessel (shaded yellow); (c,d) result after guided homotopic thinning step; (e,f) VOI after refinement (right). Images (b,d,e) on right shows zoomed sections of images (a,c,e) on left.	98

4.11	Correct detection example: (top) iso-phase image pair <i>i.e.</i> fluoroscopic image (left) with guidewire tip and the reference image at same cardiac phase; (bottom) segmented guidewire tip in fluoroscopic image and (on right) the result of matching (feature pair) and its location in the longest track.	103
4.12	VOI retrieval rate	104
4.13	Detected VOI in 6 patient sequences	105
4.14	Detected VOI in 6 patient sequences	106
4.15	Detected vessel of intervention in 10 cardiac phases of patient E. . .	108
4.16	Image quality (IQ) demonstration: normal IQ reference image (on left) and low IQ reference image (on right).	109
5.1	Markerball segmentation with semantic information based on vessel of intervention	114

List of Tables

1.1	Phases of the PCI procedure	10
3.1	Patient wise analysis of sources of variance of the performance of empty catheter segmentation algorithm	71
4.1	Variation across the dataset of 15 patient sequences with three extreme cases. The different names of VOI are: RCA, right coronary artery; LAD, left anterior descending; RPD right posterior descending artery; LCX, left circumflex artery	100
4.2	Performance of VOIDD algorithm on 12 clinical cases of 11 patients	102
4.3	Performance of VOIDD algorithm on second dataset of 3 clinical cases	107

Introduction, clinical background and motivation

Contents

1.1	Clinical context	2
1.1.1	Cardiovascular pathologies	2
1.1.2	Treatments	4
1.2	Detailed view on PCI	6
1.2.1	The Cathlab	6
1.2.1.1	Interventional system	6
1.2.1.2	ECG acquisition	8
1.2.2	Breakdown of a PCI procedure	8
1.2.2.1	Interventional tools	8
1.2.2.2	Phases of the procedure	9
1.3	Motivation: PCI procedure modeling and its applica- tions	11
1.4	A word about the dataset	12
1.5	Context of the thesis	13

This PhD in the field of image processing falls into the context of interventional cardiology, a medical specialty where procedures are performed to cure coronary heart disease, which is first cause of death worldwide. One of the most common treatment is percutaneous coronary intervention (PCI), which is minimally invasive procedure performed under the control of real time streaming of X-ray images in an interventional suite. In this work we introduce the concept of modeling of these PCI procedures to aid the patient and the clinical team. This involves understanding the different phases of the PCI procedure by the interventional machine, which can be used to optimize the X-ray dose and contrast agent. The purpose of this thesis

is to study ways to segment different interventional tools in the X-ray images and deduce semantic information from them.

In this introductory chapter, we first present the coronary heart disease and the possible treatment options, in Section 1.1. This is followed by a detailed view on the PCI procedure in Section 1.2, summarizing the clinical facts and the steps of the procedure necessary to understand the modeling of the procedure. Section 1.3 describes the concept of PCI procedure modeling, its motivation and conceivable applications. Section 1.4 is dedicated for briefly understanding the dataset used in this study. Finally, Section 1.5, presents the technical context of the thesis, the overall structure of the thesis and the contributions.

1.1 Clinical context

Ischemic heart disease, also known as coronary heart disease, and its consequences are one of the major world public health issues. According to the World Health Organization (WHO), the ischemic heart disease is the number 1 cause of death both globally and for the high income countries. According to [World Health Organization, 2017] report, 8.76 million people died of coronary artery disease in 2015 representing 15.5% of all deaths around the world. This Ph.D. thesis is made in the context of interventional cardiology, which is a key treatment option for coronary artery disease. This section aims at giving the reader a brief overview of the clinical context and introducing the terms that are commonly met in the field of interventional cardiology. We shortly introduce the cardiovascular atherosclerosis pathology in Section 1.1.1 and the associated treatment options 1.1.2.

1.1.1 Cardiovascular pathologies

Heart attack and angina are the two main consequences of the coronary artery disease. We are particularly interested in the main cause of angina or heart attack: the atherosclerosis. This disease is also called as atheromatous stenosis. It is the progressive buildup of cholesterol, fat, calcium and fibrous tissue, called *plaque*, on the inner wall of coronary arteries typically over decades (Figure 1.1). In a first stage, the vessel wall enlarges to include the presence of plaque, while maintaining a normal cross section size. Artherosclerosis is a chronic disease that remains asymptomatic for decades until this narrowing, also called stenosis, limits the flow of oxygen-rich blood to organs. This narrowing of artery lumen is usually evident in blood stream column irregularities visible on a *angiography* (Figure 1.2), an X-ray technique that enables vessel visualization.

When the stenosis is severe enough, it can have consequences on the blood supply to related part of the heart muscle, resulting in a severe chest pain due to ischemia (a lack of blood, hence a lack of oxygen supply). It is called angina pectoris (commonly known as angina). From a quantitative standpoint, a stenosis is considered significant if the narrowing obstructs more than 70% of the diameter of the artery and non-significant below 50% [Brasselet et al., 2002]. For intermediate

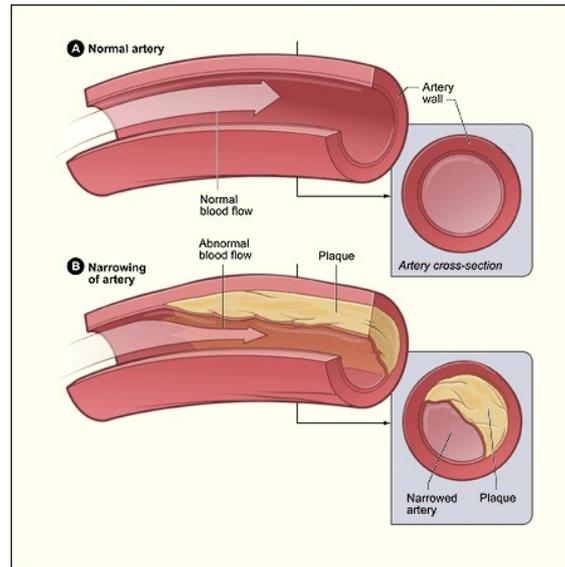


Figure 1.1: Artherosclerosis explained by [National Heart, Lung and Blood Institute, 2015*d*]. (top) Normal artery with normal blood flow; (bottom) an artery with plaque buildup. The inset images shows cross-sections of corresponding arteries.

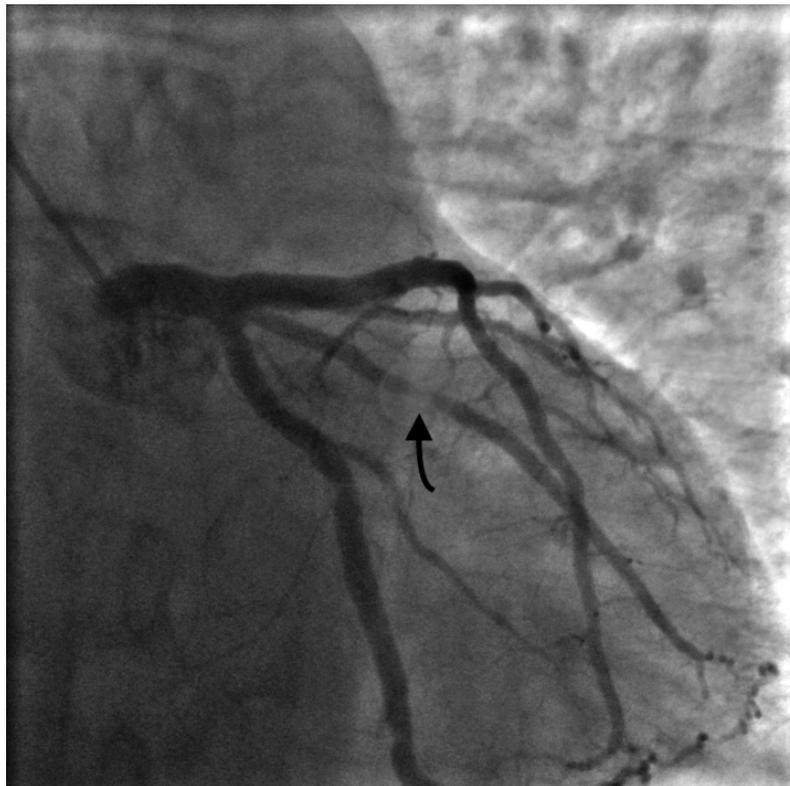


Figure 1.2: Visualization of a diseased artery under X-ray fluoroscopy (angiography) with an IGS 520 X-ray system (GE Healthcare) . The position of the stenosis is shown with an arrow.

cases, additional diagnosis and information is required to quantify the impact of the stenosis.

At a more severe stage, plaques can become unstable, rupture, and additionally promote a thrombus (blood clot) that occludes the artery. If a thrombus completely blocks an artery, a whole part of the myocardium is cut off any supply, it leads to myocardial infarction, known as heart attack. The non-irrigated cardiac cells will eventually die if the blood flow is not restored soon enough, creating collagen scars in the muscle that can induce severe life threatening clinical complications. As a thumb rule, the longer the infarct, the greater the damage.

1.1.2 Treatments

The treatment options for angina and acute myocardial infarction do share some similarities. Most of the known risk factors for atherosclerosis are: lack of physical activities, smoking, unhealthy diet, diabetes, geographical factors and high cholesterol level. Therefore the first treatment of atherosclerosis constitutes of working on these causes. This can be achieved by change in lifestyle (like exercising, no smoking and low fat/salt diet) and administration of drugs. The administered drugs are inspired by strategies to reduce the oxygen needs of the heart by reducing the heart beat rate or to dilate the coronary arteries to improve blood supply or by avoiding blood clots formation. For more severe cases, two medical procedures can be performed to treat consequences and re-vascularize properly the myocardium.

Coronary artery bypass graft (CABG): CABG is the surgical procedure performed to relieve angina and reduce the risk of death from coronary artery disease. It is an open heart surgery where arteries and/or veins from elsewhere in the patient's body are grafted to the coronary arteries to bypass the narrowing and re-orient the flow, thus improving the blood supply to the coronary circulation (Figure 1.3). Historically, it was the first developed approach. It is still recommended for the most complex cases involving multiple territories at risk.

Percutaneous Coronary Intervention (PCI): The *Percutaneous Coronary Intervention* (PCI) , or *angioplasty* is a very mature and minimally invasive procedure that consists of introducing a balloon into the clotted artery and inflating it to dilate the artery. It relies on the deployment of a stent, a fine mesh of wire, having the shape of the artery at the location of the stenosis as a permanent endoprosthesis. The main steps of these procedures, balloon angioplasty and stenting, are illustrated in Figure 1.4. We further make a detailed description of this minimally invasive procedure in next section (Sec. 1.2) since it is the procedure of interest for our work.

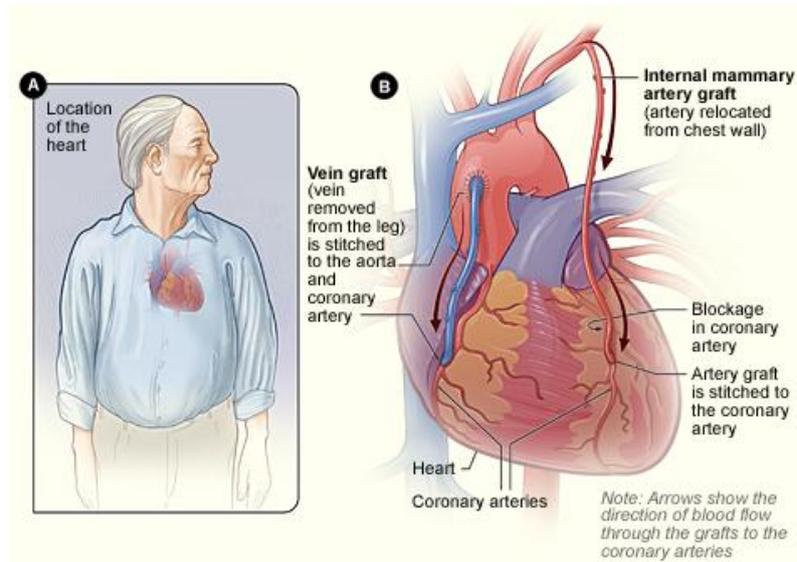


Figure 1.3: An illustration of CABG [National Heart, Lung and Blood Institute, 2015b]

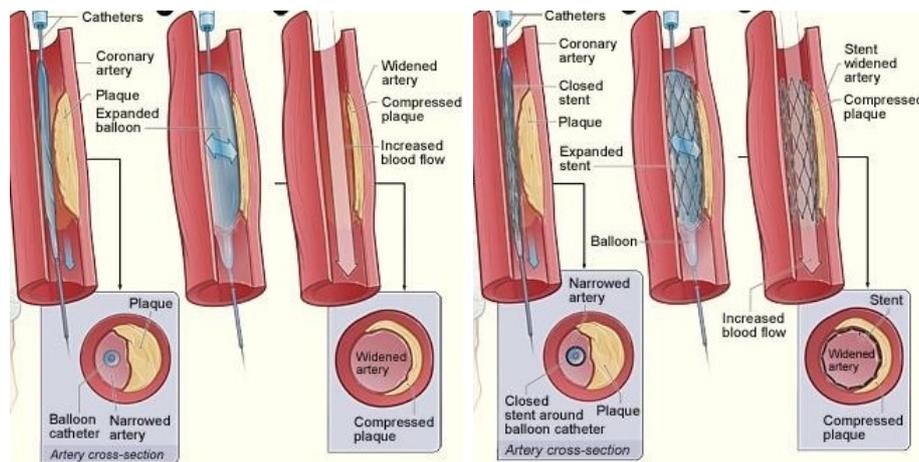


Figure 1.4: An illustration of balloon angioplasty and stenting [National Heart, Lung and Blood Institute, 2015c]

1.2 Detailed view on PCI

So far, we presented a general clinical context. We now propose to detail further the field of interventional cardiology and stenting. A very short historical perspective [Gaspard, 2017] of coronary angioplasty will enable us to understand the PCI procedure as it is performed today. A different family of methods, the angioplasty methods, has been introduced in 60' following the pioneering work of [Dotter and Judkins, 1964], that is considered as the father of Interventional Radiology and won the Nobel prize in 1978. During his first angiography, Dotter injected contrast media into the vessel to make them visible on radiographs. In 1977 Andreas Gruntzig performed first balloon coronary angioplasty (Figure 1.4), it was the birth of interventional cardiology. In order to overcome the problem of re-stenosis after balloon angioplasty, coronary (bare metal) stents were developed in mid 1980s. In 1986 Sigwart et. al. implanted the first coronary stent (Figure 1.4). The PCI procedures with stenting reduced the rate of CABG. More recently, drug eluting balloons and drug eluting stents were developed to solve the re-stenosis caused by intrastent growth of scar tissue.

1.2.1 The Cathlab

The PCI procedures take place in a dedicated interventional room called catheterization lab, or simply *cathlab*. At the center of cathlab is the imaging system in our case a system from the Innova series manufactured by GE (Interventional Guiding System: IGS 520, or IGS 530 for instance). We first detail the interventional system in the Section 1.2.1.1. In most of the cases, cathlabs are also equipped with ECG monitors, as detailed in the Section 1.2.1.2.

1.2.1.1 Interventional system

The imaging of coronary arteries (angiography) and the deployment of stent (PCI) are minimally invasive techniques monitored under X-ray fluoroscopy with the Innova system. At the command (stepping on/off the pedal) of the clinician, real-time X-ray video is produced at 7.5, 15 or 30 fps. Any part of the body can be imaged, at virtually any angulation, thanks to positioning capabilities of the system: a moving table (3 degrees of freedom in translation), on which the patient lies, and an imaging C-arm (3 degrees of freedom in rotation) that can be positioned around the patient body part. The images are displayed on the suspended monitors, right in front of the clinician. Figure 1.5 shows the Innova IGS 520 system, and the positioning of the patient, the clinician and the monitors. The system also offers the possibility to spin the C-arm around the patient to perform 3D reconstruction for instance.

Interventional systems are provided with two key acquisition modes, namely *fluoroscopy* (also called fluoro, or scopy) and *record* mode (also called cine or graphy). Fluoroscopy mode is used for maneuvering the interventional tools. In record mode, the system is set to deliver images with a quality sufficient to support the operator in his assessment of the vasculature. Record images are more contrasted and less noisy than fluoroscopic images. In fluoroscopy, the intensity of X-ray beam



Figure 1.5: Innova Interventional Guiding System, IGS-520 from GE Healthcare in cath-lab.

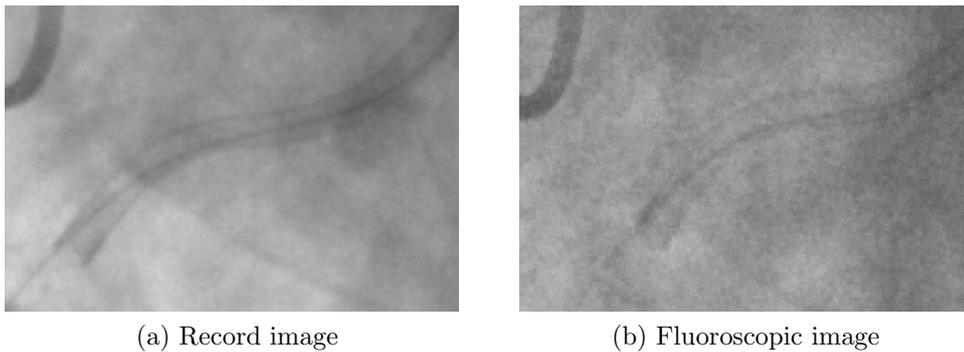


Figure 1.6: Comparison between fluoro and record images on IGS-520. All other components of the imaging situation (patient, geometry) are identical. Fluoroscopic images are noisier and the contrasts are weaker than in the record images.

and so the dose delivered are limited as per regulation. [Didier et al., 2016] reports observed radiation level and exposure time during PCI procedures in the different exposure modes. Typical number of dose rate is a ratio of 6 to 10 between the two acquisition modes which explains the significant difference in noise level and contrast. In Fig. 1.6, a record image and a fluoroscopic images taken at few seconds of interval in the same setting illustrate the difference of quality between these two imaging modes of interventional angiographic units. The appropriate usage of both and of the multiple dose customization capabilities of the system enables performing interventions according to the ALARA principle that rules the image quality/dose trade-off: Xray dose must be As Low As Reasonably Achievable.

1.2.1.2 ECG acquisition

During the PCI procedure, the patient on the table in the cathlab is connected to an ECG monitor that records the electrical activity of the heart and monitors the heart during the procedure using small electrodes that stick to the skin of the patient. These ECG monitors are also coupled with the imaging system such that the ECG signals are acquired simultaneously and are mapped temporally to the X-ray image stream.

1.2.2 Breakdown of a PCI procedure

The PCI procedures follow a set of pre-defined steps/phases. It is important to understand these different steps given the context of the thesis. A presentation of the overall procedure is provided in [National Heart, Lung and Blood Institute, 2015a], which illustrates the different phases of the PCI procedure. Several interventional tools are involved in the PCI procedure. We list few tools, visible in X-ray images, in the Section 1.2.2.1, which will help the reader to understand the different phases of the procedure detailed in Section 1.2.2.2. We refer to this breakdown in several chapters depending on the problem we are addressing.

1.2.2.1 Interventional tools

Facilities performing PCIs must have a varied inventory of coronary diagnostic/guiding catheters, coronary guide wires, angioplasty balloons, coronary stents and other treatment devices commensurate with the scope of services. These are the four tools mostly visible in X-ray images. Some details about them are as follows:

1. **Coronary catheter.** A catheter is a flexible surgical tube or instrument inserted in body cavity or vessel. The coronary catheters used during PCI are of two types: i) the diagnostic catheter, and ii) the guiding catheter. The diagnostic catheter functions as a conduit for contrast, fluids, and pressure measurement during cardiac catheterization of coronary arteries and the left ventricle. The coronary guide catheter provides support for device advancement (stents, balloons, etc.). It is the conduit for device and wire transport, a vehicle for contrast injection and takes measurements. Both these types of catheters look almost alike in X-ray images. Coronary catheters are available in a variety of shapes and sizes to fit a variety of patient anatomies. They are available in 1.5 to 3 mm inner diameter measurements.
2. **Coronary guidewire.** Coronary guidewires are designed to navigate vessels to reach a lesion or vessel segment. Once the tip of the device arrives at its destination, it acts as a guide that balloon catheters and stents can rapidly follow for easier delivery to the treatment site. The guidewire, a very thin (wire-like) object of diameter 0.35 mm has two sections namely the guidewire body and the guidewire tip. The tip being the distal section of 20 mm length, aids the navigation of guidewire. It is enough radio opaque to be seen with low-dose fluoroscopy mode of interventional systems. Guidewires with different

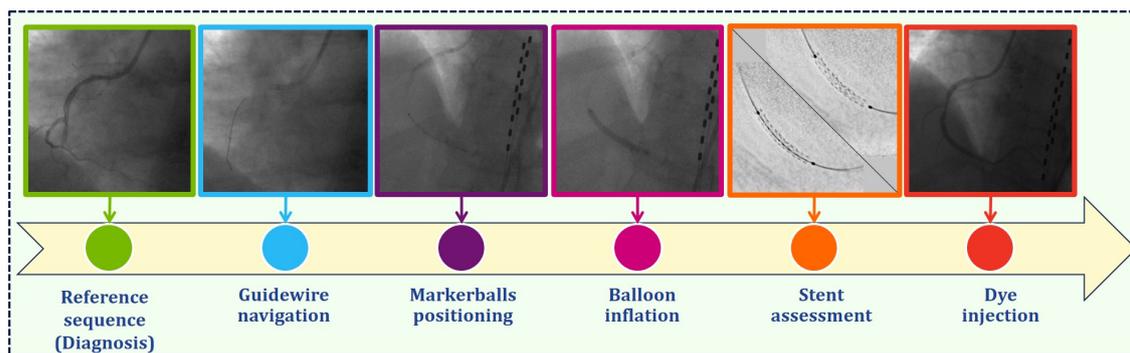


Figure 1.7: Phases of the PCI procedure

features exist and are used in specialized procedures such as treatment of chronic total occlusions.

3. **Angioplasty balloon.** An angioplasty balloon is a medical device that is inserted into a clogged artery and inflated to clear blockage and allow blood to flow. A balloon catheter is a catheter with an inflatable "angioplasty balloon" at its tip which is used during a catheterization procedure to enlarge a narrow opening or passage within the body. The balloons are provided with radiopaque markers, called *markerballs*, which helps the clinician to navigate the balloon to the location of the lesion. The deflated balloon catheter is thus positioned, then inflated to perform the necessary procedure, and deflated again in order to be removed. When a balloon catheter is used to compress plaque within a clogged coronary artery it is referred to as a plain old balloon angioplasty or POBA. Balloon catheters are also utilized in the deployment of stents during angioplasty.
4. **Coronary stent.** A stent is a wire-mesh tube placed in the coronary arteries that supply the heart, to keep the arteries open in the treatment of coronary heart disease during PCI. Stents reduce chest pain and have been shown to improve survivability in the event of an acute myocardial infarction. Balloon catheters are supplied to the cath lab with a stent premounted on the balloon. When the cardiologist inflates the balloon it expands the stent. When the cardiologist subsequently deflates the balloon the stent stays behind in the artery and pulls it back out of the patient, the stent stays (Figure 1.4).

1.2.2.2 Phases of the procedure

The different phases of the PCI procedures are shown in Fig. 1.7. As illustrated in [National Heart, Lung and Blood Institute, 2015a], firstly a needle is used to make a small hole in the blood vessel, usually the groin and the diagnostic catheter is inserted to navigate through the aorta. The radial artery, on the left or right side arm of the patient is also commonly used as access point to enter in the vasculature. The tip of the catheter is placed at the coronary ostium (entry) to produce

Table 1.1: Phases of the PCI procedure. The first column is the index of the phase, second one its description, third one is X-ray mode (F stands for fluoro and R for record) and the last one the clinical challenges.

	Phase	X-ray	Clinical challenges
1	Diagnosis	R	Assessing if lesion shall be treated; choosing appropriate sized balloon and stent
2	Guidewire navigation	F	Some lesions can be difficult to cross
3	Markerball positioning	F	Visualizing the position of the balloon versus the anatomy
4	Balloon inflation/ (stent expansion)	F or R	Controlling the pressure in balloon and counting time of inflation simultaneously
5	Stent assessment	R	Assessing is stent is correctly deployed
6	Contrast injection	R	Assessing if lesion is treated
7	If the lesion is not properly treated repeat 4	F	
8	Removal of interventional tools	F	

reference images of the pathology. This phase is referred to as *diagnostic phase* or reference sequence. The images of reference sequence, illustrating the vessels, help the clinician to decide the treatment strategy. If stenting is selected, he/she will preserve this reference image on one of his monitors to guide his gesture all along the procedure. The imaging C-arm typically stays in the same position until the end of the treatment.

Then during the *guidewire navigation phase*, a guidewire is inserted down to the distal end of the artery. The crossing of the lesion may present some difficulties depending on its characteristics. Usually, a balloon catheter is slid over the guidewire to the lesion and inflated at high pressure (10-12atm). Then a stent is mounted on a balloon and is slid down to the lesion and positioned with the help of markerballs. We term this phase as *markerballs positioning phase*. Following this, during the *balloon inflation phase*, the balloon carrying stent is inflated to deploy the stent at the lesion. Then, the balloon is removed and a stent enhancement visualization software, like StentViz (commercialized by GE), is launched. This is termed as the *stent assessment phase*. Finally, a record acquisition is performed with contrast injection to verify that the lesion is properly treated, the corresponding phase being *dye/contrast injection phase*. If necessary, a balloon is brought back and inflated again. The main phases of the procedure, the corresponding X-ray acquisition mode and the challenges met by physicians during these phases are reported in Table 1.1.

1.3 Motivation: PCI procedure modeling and its applications

In interventional cardiology, PCI procedures are performed with real time streaming of X-ray images, most of it being low dose fluoroscopic images. The clinical contribution of these procedures is very recognized: it is thus that about 175,000 stents are deployed per year in France and 2.3 million in the OECD countries [OECD, 2011]. Therefore, optimizing the cost and the quality of the result is of utmost importance for the stakeholders. In fact, the high amount of ionizing radiations and contrast agent is one of the limitation of these procedures, which impacts patients as well as the practitioner. In this context, it is often observed that the machine is used sub-optimally because it is difficult for practitioner to optimize operating points given the multiple tasks he should perform. As a result, the image quality / radiation dose / contrast agent compromise employed is not always adequate.

The design of the machine and the interactions of the imaging equipment with its user can be improved. These can be partly achieved by optimizations of the behavior of the machine according to the phases of the procedure. Getting information about phases of the procedure directly from the human operator is not acceptable from a workflow point of view. So we aim at designing a family of image processing algorithms to identify the presence of different interventional tools in the images and link this information to high-level knowledge describing the phases of the procedure and the user expectations for each of them. The understanding of these phases depends in part on a semantic analysis of the images produced throughout the procedure. We term this semantic analysis of the procedure as “*PCI procedure modeling*”.

We can relate the PCI procedure modeling to the concept of human activity recognition in computer vision. The vast research on human activity recognition has been actively investigated [Turaga et al., 2008] and translated to the domain of surgical applications [Lalys and Jannin, 2014], where it is termed as surgical process modeling. In the field of operating theater monitoring and surgical process modeling for laparoscopic and cataract surgeries, similar pioneering work has been reported [Lalys et al., 2012; Padoy et al., 2012]. In our case, such semantic information may also be used for automatic dose control.

In the cathlab, a number of events are recorded, such as the position of the imaging system, ECG information or the use of the different image acquisition modes. The most significant information is in X-ray images produced and used by the operator for guidance. It is from these images that we think we can follow the progress of the procedure. The task of PCI procedure modeling can be conceived on several levels. At the level of the individual images: by identifying the tools present in the image. At the sequence level: by identifying the intention of the operator during the sequence like what information was obtained, which tool manipulation was performed. Finally, at the level of the whole procedure: by recognizing the key phases of the procedure in the succession of recorded sequences. This strategy is based on the assumption that recognition of an interventional tool in X-ray image leads to information that a given phase of the procedure is starting or on-going.

Monitoring the interventional tools like coronary catheters, guide wire, marker balls, balloon, stent is thus necessary to obtain this information. Therefore, segmentation of these tools is a fundamental brick in such procedure modeling. It is then required to obtain semantic information about the regions of interest, control points and phases using the segmented tools. The appearance of these different devices may vary a lot in the projected X-ray images. The numerous factors contributing to this variation may be the size and shape of the object, radio-opacity of the material and the orientation of the tool in the imaging plane. These tools are contrasted and relatively easy to segment. Others are more challenging and we aim to develop collaborative segmentation scheme. Tracking these tools along the sequence and obtaining semantic information is challenging given the complex motions apparent because of heart beat, breathing and the navigation of the interventional tools.

PCI procedure modeling can help to improve the interaction of the clinician with the imaging equipment. Some of the applications that we can envision can be smart dose control, as patented in [Riddell et al., 2017]. The labelled sequences according to recognized phases can be used to generate automatic electronic medical records. A system was presented for automatic generation of electronic medical record [Agarwal et al., 2007], which contains information about administration of medicines and the occurrence of medically significant events. In case of PCI procedures, clinicians have mentioned importance of structured reporting and comprehensive documentation of indications for PCI procedures [Nallamothe et al., 2014; Sanborn et al., 2014]. Besides, identification of key events can be used to automatically launch softwares like StentViz, thus improving the user machine interaction. Another application of PCI procedure modeling, can be prediction of remaining time which would aid the clinical teams for smoother clinical workflow between two interventions. Similar application has been presented in the field of surgical procedure modeling [Aksamentov et al., 2017].

1.4 A word about the dataset

During the PCI procedures, in general, the record images are automatically stored on the system and are used for diagnosis and documentation of the stages of the treatment. Whereas the fluoroscopic images are not systematically stored. In some cases, like ours, based on an agreement between the clinical institutes and research labs, the fluoroscopic images throughout the procedure can be stored and then retrospectively used for research purposes. Thus, our dataset consists of fluoroscopic and record images along with the mapped ECG data. Each image is also associated with additional system information like the position of the C-arm gantry, the table position, collimation parameters, and time stamp for the frame. Though the dataset used in this work is from a single clinical site, there is a lot of variation because of different patients, anatomy, SNR, dose etc. We further mention details on data used in different parts of work in the corresponding chapters.

1.5 Context of the thesis

In the previous paragraph we presented PCI procedures which are minimally invasive treatment for coronary heart disease and the concept of PCI procedure modeling we introduce in this work. In order to perform continuous monitoring of the procedure to reach the goal of PCI procedure modeling, one of the important task is to detect different interventional tools in the X-ray image sequences and deduce semantic information from them and various available information. Thus, it is necessary to detect, recognize and delineate automatically (segment) the interventional tools in record as well as fluoroscopic images. However, the segmentation task is difficult due to low contrast in fluoroscopic images. Given the different nature of the tools to be detected (shape / contrast / size), the methods to detect should take advantage of multi-scale descriptions of images to construct appropriate hierarchical image representations and detect the structures of interest.

Segmentation of different interventional tools has been studied vastly in literature for aiding or driving different applications: motion compensation, navigation, stent visualization, 3D/2D registration, guidewire enhancement, denoising, 3D reconstruction of stents and closure devices. Overall we observe that dedicated methods were proposed in the literature for the detection and the segmentation of specific interventional tools. However, these methods cannot be generalized for segmentation of other tools. We propose methods based on graph based connected filters for segmenting these interventional tools. We choose a unique framework of mathematical morphology based on component trees, a tree based shape-space, for segmenting different interventional tools. The goal of Chapter 2 is to have a brief literature review of the methods used for segmentation of interventional tools and to have an overview of the background and theory of mathematical morphology based connected operators which are the core segmentation methods of this thesis. We also review some registration methods in this chapter.

We split this work in two parts, namely: 1) empty catheter segmentation and, 2) vessel of intervention detection by tracking guidewire tip. In Chapter 3, we present a method for the segmentation of empty coronary catheter in X-ray images. The coronary catheter, being a commonly visible landmark, its segmentation is an important and a difficult brick for PCI procedure modeling. In a number of clinical situations, the catheter is empty and appears as a low contrasted structure with two parallel and partially disconnected edges. To segment it, we work on the level-set scale-space of image, the component tree called as min tree, to extract curve blobs. We then propose a novel structural scale-space, a hierarchy built on these curve blobs. The deep connected components, which are the clusters of curve blobs on this hierarchy are analyzed to detect empty catheter. This work was presented in conferences [Bacchuwar et al., 2016] and was published in journal [Bacchuwar et al., 2017a].

In Chapter 4, we present a fully automatic tracking algorithm, called vessel of intervention dynamic detection (VOIDD), to detect the vessel of intervention (VOI), which is treated during the procedure, by combining information from a reference image with contrast injected vasculature and fluoroscopic images acquired

during guidewire tip navigation phase where the VOI is not visible. This work is preliminary step towards PCI procedure modeling and can be used to automatize preparation of enhanced images and also to optimize the image acquisition to reduce the use of ionizing radiation or amount of contrast media. This work was presented in a conference [[Bacchuwar et al., 2017b](#)] and the corresponding journal article is under review.

Both the chapters on the contributions of this PhD thesis are organized as an independent publication (abstract - introduction - method - result - discussion - conclusion) and can be read quite independently from the other ones. Readers who are interested in mathematical morphology and graph based connected filters can read [Chapter 2](#) independently.

Related works

Contents

2.1	Segmentation and detection of interventional tools in X-ray images	16
2.2	Graph based connected filtering	18
2.2.1	General overview	18
2.2.2	Graphs and connectivity for image analysis	19
2.2.3	Connected operators	23
2.2.3.1	Mathematical morphology operators on binary and grayscale images	23
2.2.3.2	Definition of connected operators	24
2.2.3.3	Example of connected operators: attribute filters	25
2.2.4	Connectivity in weighted graphs	28
2.2.4.1	Weighted graphs	28
2.2.4.2	Stacks of graphs	29
2.2.4.3	Stacks induced by a vertex-weighted graph	30
2.2.4.4	Stacks induced by an edge-weighted graph	31
2.2.5	Connected partitions and hierarchies	32
2.2.6	Component trees: tree based shape space	36
2.2.6.1	Min/Max tree.	37
2.2.6.2	Quasi-flat zones tree	40
2.2.6.3	Algorithms for component tree construction	42
2.2.7	Filtering on component trees	43
2.2.7.1	Increasing attributes based filtering	44
2.2.7.2	Non-increasing attributes based filtering	44
2.2.7.3	Filtering with shaping	45
2.2.8	Challenges with connected filters: second order connectivity	47
2.3	Registration	48

2.1 Segmentation and detection of interventional tools in X-ray images

The task of segmentation and detection of interventional tools in X-ray fluoroscopic images have been studied in details by various teams. They mostly aimed at real time or near real time detection of interventional tools like marker balls, guidewires, stents, IVUS probes, and catheters. Most of these methods were developed for aiding or driving different applications: motion compensation, navigation, stent visualization, 3D/2D registration, guidewire enhancement, denoising, 3D reconstruction of stents and closure devices. The appearance of these different devices may vary a lot in the projected X-ray images. The numerous factors contributing to this variation in appearance may be the size and shape of the object, radio-opacity of the material and the orientation of the tool in the imaging plane.

Many methods in the literature use the marker balls to detect the stent regions. The proposed applications are either 2D digital stent enhancement/visualization [Florent et al., 2008; Bismuth et al., 2011] or 3D stent reconstruction [Schoonenberg et al., 2008, 2009]. [Schoonenberg et al., 2009] describes the marker balls as blob like structures. They use blob detection with automatic scale selection proposed by [Lindeberg, 1993], thus obtaining number of candidate markers. The tracked marker couple is further used to perform the task of motion compensation or reconstruction of coronary segments. The authors of [Bismuth et al., 2011] proposed marker ball segmentation based on local minima selection in an intensity range. The numerous detected candidates for marker balls were then tracked to detect the marker balls couple, which was followed by guidewire segmentation between the marker balls. Finally, the stent is visualized by combining images after registering them in a non-linear fashion to a reference position (image). [Chen et al., 2012; Wang et al., 2012] accomplish the task of marker ball detection by training a detector with manually annotated samples and probabilistic boosting tree classifier. In these two methods, the tracking is based on Vertibi algorithm to find the marker ball couple in globally optimal manner.

The various guidewire enhancement techniques include edge-detection methods [Palti-Wasserman et al., 1997], methods based on vesselness measure [Frangi et al., 1998; Heibela et al., 2009] and Hessian eigenvalues [Baert et al., 2003] and phase congruency [Slabaugh et al., 2007]. The various guidewire detection methods were reviewed, compared and classified into three families by [Bismuth et al., 2009]. The application of clinical interest for this study was respiratory motion compensation for 2D/3D roadmapping. They claim that the Hessian based methods are most robust to strong curvature of the guidewires and that techniques based on family of highly anisotropic rotated filters [Kunz and Schweiger, 2005] are most suited for detection of low CNR and low curvature guidewires. They also demonstrate that steerable filter approach, based on the work of [Jacob and Unser, 2004], had less interesting detection capabilities and appeared to be most computationally expensive among the three families.

Among the another set of methods based on machine learning, [Barbu et al., 2007] presented guidewire localization and enhancement in cardiac fluoroscopic im-

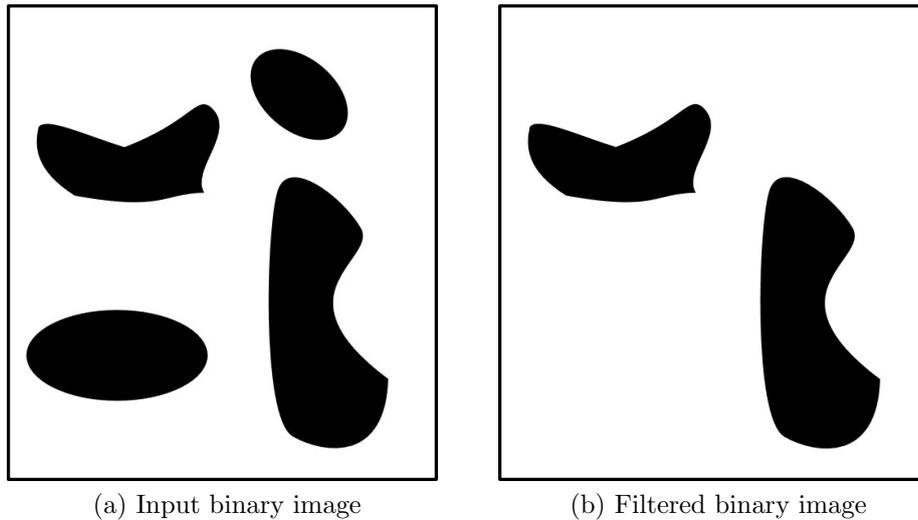


Figure 2.1: An example of binary connected operator.

ages using hierarchical data organization guided by classification. As a first step, they obtain parts of curve using a low-level segment detector. They use marginal space learning based hierarchical model of curves to model complex free-form curves. This method is purely learning based and is neither invariant to guidewire pose parameters nor it can handle well miss-detection. In order to overcome these shortcomings [Honnorat et al., 2010b,a] proposed more robust guidewire segmentation method based on detection using boosting, mid-level grouping scheme based on clustering and a complete reconstruction through the minimization of global criterion. Other method for guidewire enhancement include more general curvilinear structure enhancement technique based on polygonal path image as proposed by [Bismuth et al., 2012]. This method is based on selection of best fitting curve at each pixel and has ability to control the smoothness and length of the curvilinear structures to be analyzed. This method unifies in a single framework local, semi-local and global curvilinear structure analysis with computational efficiency.

Catheters of different types are also another important set of tools in interventional procedures which serve for different purposes. We have listed several works on segmentation and detection of various catheters in Section 3.2.2 as they are more relevant to the empty catheter segmentation problem we address in this work. Overall we observe that dedicated methods were proposed in literature for detection and segmentation of specific interventional tools. However, these methods cannot be generalized for segmentation of other tools.

PCI procedure modeling requires segmentation and detection of various interventional tools which appear throughout the procedure. We propose methods based on graph based connected filters for segmenting these interventional tools.

2.2 Graph based connected filtering

This section presents the background of the core segmentation methods used in this thesis. We present the basic definitions and notions of graphs in mathematical morphology in the Section 2.2.2. We review the mathematical morphological connected operators in graph based framework in Section 2.2.3. Following this we present the notions of connectivity in weighted graphs in Section 2.2.4 and the concept of connected partitions and hierarchies in Section 2.2.5. We then introduce the concept of component trees as tree-based shape space for tree-based hierarchical representations of images (graphs) and filtering on these shape spaces in Section 2.2.6 and Section 2.2.7. We then discuss the challenges for using graph based connected filters and some approaches to solve them in Section 2.2.8.

2.2.1 General overview

Traditionally, digital images are represented as rectangular grid of pixels and digital videos are seen as continuous flow of digital images. New multimedia applications like indexing (*e.g.* PCI procedure modeling presented in this thesis), require a representation that is closer to observed and human perceived physical reality. The tasks in indexing applications like creation of table of contents rely on the knowledge of spatio-temporal entities describing when and where the indexed elements can be observed. In these applications, the notion of regions is central. If we try to distinguish between the notion of regions and objects: an object is a set of regions that forms a semantic entity. Most of the signal processing tools, in particular classical filtering techniques, are inappropriate for region-based representation and processing as they depend on classical pixel-oriented signal representation where the notion of region is absent. In literature, early examples of region-based processing can be found in the field of segmentation.

Connected filters [Serra and Salembier, 1993; Salembier and Serra, 1995; Salembier et al., 1998; Salembier and Wilkinson, 2009] is a set morphological filtering tools which has received much attention being at the frontier between segmentation and filtering [Gatica-Perez et al., 2001]. Connected filters have the property of simplifying the image while preserving contour information and have proved useful in a large number of applications, including image filtering, segmentation, pattern recognition, and multiresolution decomposition. *Connected filters* are filtering tools which can be defined to be acting on sets or graphs. They focus on the notion of connected components, *i.e.*, maximal set of vertices in which a path exists between any two vertices. The extension of connected operators to different types of graphs (vertex or edge weighted) leads to definition of several hierarchical representation: the component tree [Salembier et al., 1998], the binary partition tree [Salembier and Garrido, 2000], the tree of shapes [Monasse and Guichard, 2000b] or the quasi-flat zones tree [Cousty, Najman and Perret, 2013]. They are higher level of abstraction for understanding [Chen et al., 2000], classifying [Urbach et al., 2007], filtering [Salembier et al., 1998] or segmenting [Jones, 1999; Passat et al., 2011] regions (and eventually objects) in images. In this PhD, the connected filters-based concepts which are proved in literature are used for images and more complex spaces

than images. As mentioned by [Perret et al., 2015], a general definition scheme for connected filter consist of four steps:

1. construct the image hierarchical representation;
2. compute attributes at each node of the representation;
3. select relevant nodes according to these attributes;
4. produce a filtered image or segmentation map.

A connected component appearing in the hierarchical representation is called as *node*. Based on the above four steps, the task of designing connected operators to perform segmentation, classification or filtering has two major issues, noted:

1. making structures of interest emerge in the hierarchical representation;
2. distinguishing structures of interest in this hierarchy.

The first issue has been discussed in the literature through the definition of second generation connectivity [Ronse, 1998; Ouzounis and Wilkinson, 2007a; Passat and Naegel, 2011], constrained connectivity [Soille, 2008], and hyper-connectivity [Ouzounis and Wilkinson, 2011]. Second generation connectivity will be discussed in Section 2.2.8. For the first issue, our contribution of deep connected components in multi-level scale space will be described in Chapter 3. The second issue, *i.e.* selecting relevant nodes of hierarchy refers to steps 2 and 3 of the definition scheme of connected operators:

- defining attributes that provide a suitable feature space to be able to characterize relevant nodes; and
- defining robust and accurate node selection process.

The task of defining attributes refers to designing attribute functions described in the Section 2.2.3.3. Significant effort has also been made to propose node selection processes which have evolved from global thresholding [Jones, 1999], to energy-minimization methods [Guigues et al., 2006; Serra, 2012], to connected filtering in feature space [Xu et al., 2016] and learning based classification [Caldairou et al., 2009]. More details on attributes is provided in tree based image filtering strategies described in Section 2.2.7. The notion of *connectivity* is required to define a region mentioned above in this section. The objective of the next section is to introduce the elementary notions of the graph theory which provides, in particular, the tools to understand the concept of connectivity in the discrete sets.

2.2.2 Graphs and connectivity for image analysis

Graphs are effective representation applied vastly in image analysis and processing [Lézoray and Grady, 2012]. The usefulness of graphs in mathematical morphology has been recognized by [Vincent, 1989], and later [Bertrand, 2005; Cousty et al.,

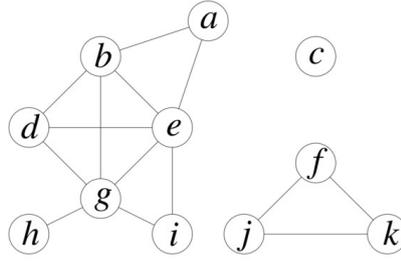


Figure 2.2: A graph

2009; Heijmans and Vincent, 1999; Lom enie and Stamon, 2008; Meyer and Angulo, 2007; Meyer and Lerallut, 2007; Stawiaski and Meyer, 2009; Ta et al., 2008; Couprie et al., 2011] extended various tools from mathematical morphology in graphs. Further, the overview of advantages of graphs for mathematical morphology was described with least possible mathematical jargon by [Najman and Cousty, 2014]. Graphs are more abstract representation than images with adjacency relationship and less specific than connectivity. In mathematical morphology, set connectivity is explained through more global and axiomatic definition of *connection*, first proposed by [Serra, 1988] (Ch. 2, p. 51).

Definition 2.2.1 (Graphs). A *graph* is a pair $\mathcal{G} = (V, E)$ composed of a non-empty finite set V and of a set E of unordered pairs of elements of V , *i.e.*, E is a subset of $V \otimes V = \{\{x, y\} \subseteq V \mid x \neq y\}$. Each element of V is called *vertex* (of \mathcal{G}), and each element of E is called an *edge* (of \mathcal{G}). If $V \neq \emptyset$, we say \mathcal{G} is *non-empty*. The vertex and edge sets of a graph \mathcal{G}' are denoted by $V(\mathcal{G}')$ and $E(\mathcal{G}')$ respectively.

Let \mathcal{G} be a graph. Similarly, we consider a function $\Gamma_{\mathcal{G}}^*$, of $V(\mathcal{G})$ in the set of parts of $V(\mathcal{G})$, which associates to each vertex x of \mathcal{G} to set of its neighbors: $\Gamma_{\mathcal{G}}^*(x) = \{y \in V(\mathcal{G}) \mid \{x, y\} \in E(\mathcal{G})\}$. Let x and y be two vertices of \mathcal{G} . If $y \in \Gamma_{\mathcal{G}}^*(x)$, we say x and y are adjacent for \mathcal{G} .

The Figure 2.2 below shows an example of a graph \mathcal{G} . The vertices are represented by the circles and the edges by the segments joining two vertices. The set of vertices of \mathcal{G} is $V(\mathcal{G}) = \{a, b, c, d, e, f, g, h, i, j, k\}$; the set $E(\mathcal{G})$ of edges of \mathcal{G} is: $\{\{a, b\}, \{a, e\}, \{b, d\}, \{b, e\}, \{b, g\}, \{d, e\}, \{d, g\}, \{e, g\}, \{e, i\}, \{g, h\}, \{g, i\}, \{f, j\}, \{f, k\}, \{j, k\}\}$. The vertex a is adjacent to vertex b and $\Gamma_{\mathcal{G}}^*(a) = \{b, e\}$.

Let i and j be two integers, *i.e.*, $i \in \mathbb{Z}, j \in \mathbb{Z}$, we denote by $[i, j]$ the set $\{k \in \mathbb{Z} \mid k \geq i \text{ and } k \leq j\}$.

Definition 2.2.2 (Paths and connectivity). Let \mathcal{G} be a graph. Let $\pi = \langle x_0, \dots, x_{\ell} \rangle$ be an ordered sequence of vertices of \mathcal{G} , we say that π is a *path from x_0 to x_{ℓ} in \mathcal{G}* if for any $i \in [1, \ell]$, the vertex x_i is adjacent to x_{i-1} , *i.e.*, $(x_i, x_{i-1}) \in E(\mathcal{G})$. The graph \mathcal{G} is said to be *connected* if for any pair of vertices x and y of \mathcal{G} , there exists a path from x to y .

In the graph \mathcal{G} of Figure 2.2, the sequence $\langle a, b, e, g, h \rangle$ is a path: the vertices a and h are connected. On the other hand, vertices b and c are not connected. Indeed

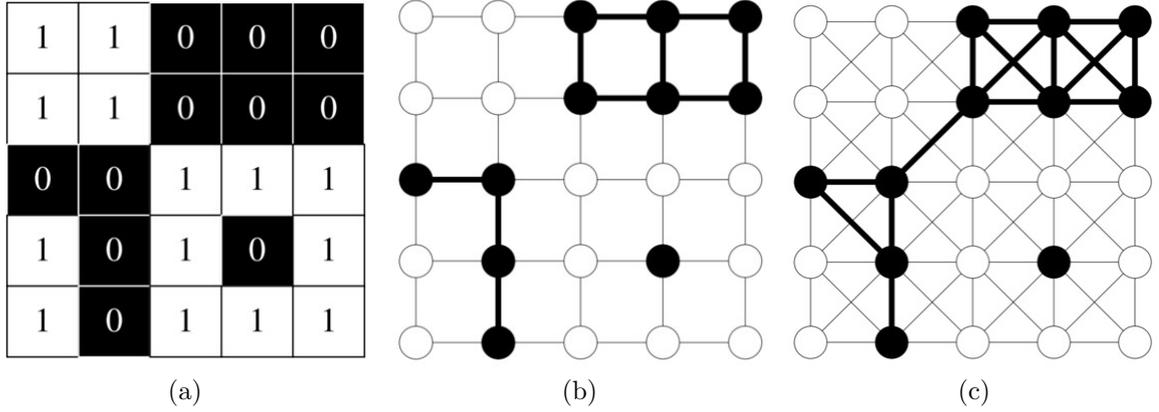


Figure 2.3: Adjacency relations for image analysis: (a) a two-dimensional image; (b) graph induced by 4-adjacency associated with the image (a); (c) graph induced by 8-adjacency associated with the image (a)

in \mathcal{G} , there is no path from b to c . The graph \mathcal{G} is not connected. On the other hand, the graph $\mathcal{G}_* = (\{f, j, k\}, \{\{f, j\}, \{f, k\}, \{j, k\}\})$ is connected. Note also that \mathcal{G} can be broken down into three connected graphs: these are the connected components of \mathcal{G} .

Let \mathcal{G} be a graph, a *subgraph* of \mathcal{G} is a graph \mathcal{G}' such that $V(\mathcal{G}') \subseteq V(\mathcal{G})$ and $E(\mathcal{G}') \subseteq E(\mathcal{G})$. If \mathcal{G}_* is subgraph of \mathcal{G} , we write $\mathcal{G}_* \sqsubseteq \mathcal{G}$.

Definition 2.2.3 (Connected components of a graph). Let \mathcal{G} be a graph and let \mathcal{G}' be a subgraph of \mathcal{G} . We say that \mathcal{G}' is *connected component* of \mathcal{G} , or simply a *component* of \mathcal{G} , if \mathcal{G}' is connected and maximal for this property, *i.e.*, for any connected graph \mathcal{G}'' such that $\mathcal{G}' \sqsubseteq \mathcal{G}'' \sqsubseteq \mathcal{G}$, we have $\mathcal{G}' = \mathcal{G}''$.

We can note that any two distinct connected components \mathcal{G}' and \mathcal{G}'' of a graph \mathcal{G} are necessarily disjoint: $V(\mathcal{G}') \cup V(\mathcal{G}'') = \emptyset$ and $E(\mathcal{G}') \cup E(\mathcal{G}'') = \emptyset$.

The graph $\mathcal{G}_i = (\{f, j\}, \{\{f, j\}\})$ (Figure 2.2) is a subgraph of \mathcal{G} but it is not a connected component of \mathcal{G} . In fact, the graph $\mathcal{G}_* = (\{f, j, k\}, \{\{f, j\}, \{f, k\}, \{j, k\}\})$ is connected and \mathcal{G}_i is a subgraph of \mathcal{G}_* ; \mathcal{G}_* , on the other hand, is a connected component of \mathcal{G} .

Usual adjacency relation for 2D image analysis. An image F is often defined as a function of set \mathcal{D} , called as *domain of the image*, into a set \mathbb{K} . The domain of a two-dimensional image is in general a “rectangular” subset of \mathbb{Z}^2 : the elements of \mathcal{D} are called pixels. The image F is often with integer values, that is to say $\mathbb{K} \subseteq \mathbb{Z}$. In image analysis, there are two fundamental adjacency relations defined on \mathbb{Z}^2 which makes it possible to structure the pixels of a two-dimensional image with a graph. The *4-adjacency relation*, denoted as E_4 , is defined by: $\forall x, y \in \mathbb{Z}^2, \{x, y\} \in E_4$ if and only if $|x_1 - y_1| + |x_2 - y_2| = 1$, where $x = (x_1, x_2)$ and $y = (y_1, y_2)$. The *8-adjacency relation*, denoted by E_8 , is defined as: $\forall x, y \in \mathbb{Z}^2, \{x, y\} \in E_8$ if and only if $\max |x_1 - y_1| + |x_2 - y_2| = 1$.

The Figure 2.3 illustrates the notions introduced in this paragraph. Figure (a) shows a two-dimensional image. Figures (b) and (c) show the graphs associated with (a) when the 4- and the 8-adjacency relations are considered, respectively.

Intuitively, when an image domain is structured with a graph as described in the previous paragraph, we can see a binary image as a subset of this domain, that is a subset of the vertex set of the considered graph. The following definition allows to infer a notion of a connected component of a binary image from the definitions on graphs seen in this section.

Subgraphs induced by sets of vertices and edges. Let \mathcal{G} be a graph. Let \mathcal{V} (resp. \mathcal{E}) be the set of all subsets of vertices (resp. edges) of \mathcal{G} . Let P be a subset of vertices of \mathcal{G} , *i.e.*, $P \in \mathcal{V}$. The *edge set induced by P* , denoted by $\epsilon(P)$, is the set of all edges of \mathcal{G} having their two vertices in P :

$$\epsilon(P) = \{\{x, y\} \in E \mid x \in P, y \in P\} . \quad (2.1)$$

Using such operator ϵ , the usual notion of a subgraph induced by a set of vertices can be easily recovered. More precisely, the *subgraph of \mathcal{G} induced by P* , denoted by \mathcal{G}_P , is the graph whose vertex set is P and whose edge set is the edge set induced by P : $\mathcal{G}_P = (P, \epsilon(P))$. Hence, the graph induced by P is the graph whose set of vertices is P and whose set of edges is composed of the edges of \mathcal{G} having their two vertices in P .

Let A be a subset of edges of \mathcal{G} , *i.e.*, $A \in \mathcal{E}$. The *vertex set induced by A* , denoted by $\delta(A)$, is the set of all vertices of \mathcal{G} which belong to an edge in A :

$$\delta(A) = \{x \in V \mid \exists u \in A, x \in u\} . \quad (2.2)$$

The subgraph of \mathcal{G} induced by A , denoted by \mathcal{G}_A , is the graph whose set of edges is A and whose vertex set is $\delta(A)$: $\mathcal{G}_A = (\delta(A), A)$. In other words, the vertex set of the graph \mathcal{G}_A induced by A is composed of all vertices which appear in an edge of A and its edge set is A . Hereafter, by abuse of notations, the subgraph of \mathcal{G} induced by the edge set A is simply denoted by A and the one induced by the vertex-set P is simply denoted by P . More information on the mathematical morphology properties of the operators ϵ and δ are detailed in article [Cousty, Najman, Dias and Serra, 2013] where it is shown in particular that δ is a dilation and ϵ is the adjunct erosion of δ .

Connected components of set of vertices and of edges. Let \mathcal{G} be a graph and let P be a subset of $V(\mathcal{G})$ (resp. a subset of $E(\mathcal{G})$). We say that P is connected whenever the subgraph of \mathcal{G} induced by P is connected and if C is a subset of P such that the subgraph of \mathcal{G} induced by C is a connected component of \mathcal{G}_P , we say that C is a connected component of P . In other words, a connected component C of P is a connected subset of P that is maximal for this property.

In Figure 2.2, $(\{g, h, i\}, \{\{g, h\}, \{g, i\}\})$ is the graph induced by the set of vertices $\{g, h, i\}$ and by the set of edges $\{\{g, h\}, \{g, i\}\}$. We can notice that there are subgraphs of \mathcal{G} that are not induced by any set of vertices. For example, in Figure 2.2, there is no subset of vertices that induces $\mathcal{G}_* = (\{a, b, e\}, \{\{a, b\}, \{a, e\}\})$. The subgraph \mathcal{G}_* is induced by the set of edges $\{\{a, b\}, \{a, e\}\}$; and the subgraph

induced by $\{a, b, e\}$ is the graph $(\{a, b, e\}, \{\{a, b\}, \{a, e\}, \{b, e\}\})$. In Figure 2.3b, (the subgraph induced by) the set of black vertices comprises of three connected components whereas in Figure 2.3c, it comprises of only two connected components.

As said previously, a binary image X is a subset of an image domain. If the image domain D is equipped with an adjacency relation E then the pair $\mathcal{G} = (D, E)$ is a graph and the binary image X is simply a subset of the vertices of \mathcal{G} . Within this setting, the notion of a connected component of the binary image X is well defined. For instance, in Figure 2.3, for a given binary image the connected components are the three (resp. two) black regions in Figure 2.3b (resp. Figure 2.3c) depending on the adjacency relationship.

A subset of vertices of graph \mathcal{G} is connected if the corresponding induced graph is connected. The notion of connectivity in the vertices of a graph can be realised with the notion of connection mentioned in the beginning of this section. As compared to set connectivity proposed by Serra [Serra, 1988], the notion of connectivity in graph theory is generated by an adjacency relation that can be described as local in the sense that it involves only pairs of vertices. Note also that the notions related to connected components seen in this section have been extended to directed graphs in [Perret et al., 2015], allowing, in particular, to handle non-symmetric relations between image pixels.

Important notation. In the sequel of this manuscript, the symbol \mathcal{G} denote a connected graph. To shorten the notations, the vertex and edge sets of \mathcal{G} are denoted by V and E respectively instead of $V(\mathcal{G})$ and $E(\mathcal{G})$.

2.2.3 Connected operators

In mathematical morphology, connected operators [Serra and Salembier, 1993; Salembier and Serra, 1995; Salembier et al., 1998; Salembier and Wilkinson, 2009] are filtering tools which can be defined to be acting on sets or can be extended to act on grayscale images. Connected operators are based on the notion of connected components, *i.e.* maximal sets of points that may be connected by a path in the set (see Section 2.2.2). In this section, we first recall the definitions of mathematical morphology operators and filters. Then, we introduce connected operators following the flaw of [Serra and Salembier, 1993]. Finally, we close the section with some basic examples of connected filters.

2.2.3.1 Mathematical morphology operators on binary and grayscale images

Let S be any set, in mathematical morphology, an (*binary-*) *operator on S* , also called a set-operator on S , is any map from $\mathcal{P}(S)$ in $\mathcal{P}(S)$, where $\mathcal{P}(S)$ denotes the set of all subsets of S . Any map f from S into the set \mathbb{R}^+ of positive real values, is called an (*grayscale*) *image on S* and we denote by $\mathcal{F}(S)$ the set of all images on S . A (*grayscale-*) *operator on S* is any map from $\mathcal{F}(S)$ in $\mathcal{F}(S)$. In other words, a binary operator Ψ maps to any subset X of S another subset $\Psi(X)$ of S and any grayscale operator ψ maps to any image f on S another image $\psi(S)$ on S .

Mathematical morphology is particularly interested in families of operators satisfying some algebraic properties. Before detailing further the applications of mathematical morphology operators, let us review these different properties. Let S be a set and Ψ be any binary-operator on S .

- The operator Ψ is *increasing* if, for any two subsets X and Y of S , we have $X \subseteq Y \Rightarrow \Psi(X) \subseteq \Psi(Y)$.
- The operator Ψ is *extensive* if, for any subset X of S , we have $X \subseteq \Psi(X)$;
- The operator Ψ is *anti-extensive* if, for any subset X of S , we have $\Psi(X) \subseteq X$;
- The operator Ψ is *idempotent* if, for any subset X of S , we have $\Psi(X) = \Psi(\Psi(X))$;
- The operator Ψ is *self-dual* if, for any subset X of S , $\Psi(X) = S \setminus \Psi(S \setminus X)$.

Similarly, for a grayscale operator ψ working on grayscale images the above mentioned properties are defined by replacing \subseteq by \leq and the complement operation ($X \rightarrow S \setminus X$) by the opposite operation $f \rightarrow -f$. Accordingly, a grayscale operator ψ is self-dual whenever it is able to process dark and bright image components symmetrically, *i.e.* $\psi(f) = -\psi(-f)$. A mathematical morphology operator is called a *filter* whenever it is both increasing and idempotent [Serra, 1988]; a filter which is extensive is called a *closing* and a filter which is anti-extensive is called an *opening*.

2.2.3.2 Definition of connected operators

We now present the definition of connected operators as provided by [Serra and Salembier, 1993]. Intuitively, in the case of a binary image X defined on a domain D (*i.e.*, X is a subset of D), a connected operator is a binary operator which can only remove from X connected components of X or add to X connected components of the complement of X . When the image domain $D = V$ is structured as a graph, the notion of a connected component of X reduces to a maximal subset of vertices included in X in which a path exist between any two points. In order to deal with the set of connected components of a subset X of V , in the following, we denote by $\mathbb{C}(X)$ the set of all connected components of X .

Definition 2.2.4 (Connected operator). A *connected (binary) operator* is a binary operator Ψ on V such that for any subset X of V , the set of symmetrical difference $X \Delta \Psi(X)$ is exclusively composed of connected components of X or of its complement $X^C = V \setminus X$, *i.e.*, there exist a subset A of $\mathbb{C}(X)$ and a subset B of $\mathbb{C}(X^C)$ such that $\bigcup\{A \cup B\} = (X \setminus \Psi(X)) \cup (\Psi(X) \setminus X)$.

Intuitively, when speaking of a binary image, this means that the connected operators act only by preserving or removing connected components of the foreground and of the background of the image. Figure 2.1 illustrates the connected operator that removes the “oval connected components” of binary images.

In order to extend this definition to grayscale operators, [Serra and Salembier, 1993] considers the set of flat zones of an image, a grayscale connected operator

being an operator which can only “enlarge” the flat zones. Let us now recall precise definition for these notions.

These connected operators filter the grayscale images by merging its elementary regions called flat zones [Salembier and Serra, 1995].

Definition 2.2.5 (Flat zone). Let f be a grayscale image on V . A *flat zone* of f is a maximal connected subset X of V where f is constant, *i.e.*, (i) X is connected, (ii) $f(x) = f(y)$ for any two x and y in X , and (iii) for any subset Y of V , Y satisfies both (i) and (ii) implies $Y = X$.

An illustration of flat zone can be seen in Figure 2.10a, where each region denoted by a separate gray-value denotes a flat zone of the image. It can be seen, as demonstrated by [Serra and Salembier, 1993], that the set of flat zones of a grayscale image f is a partition of the image domain.

Definition 2.2.6 (Connected grayscale operator). An grayscale operator ψ is said connected if, for any image f on V , any flat zone of f is included in a flat zone of $\psi(f)$.

This implies that the only allowed operation is the deletion of connected components, which is merging of flat zones. Therefore, the regions of the output partitions are created by the union of regions of the input partition, thus ensuring that they neither create new contours nor change the location or shape of contours in the input image. The connected operators are considered as filtering tools in the sense that they transform input grayscale image into filtered grayscale image. Owing to the contour-preservation properties, connected operators are capable of both low-level filtering and higher-level object recognition. An example of such low-level filtering task is shown in Figure 2.4, where connected operator is compared with other classical filters like low-pass (mean) filter, median filter, Gaussian filter and other morphological operators like opening with structuring element.

Thanks to these properties of connected operators, they have been used for filtering [Salembier et al., 1998], segmentation [Jones, 1999; Passat et al., 2011], classification [Urbach et al., 2007], and registration [Mattes et al., 1999]. Connected operators are claimed to bridge the gap between segmentation and classical filtering [Gatica-Perez et al., 2001].

2.2.3.3 Example of connected operators: attribute filters

As seen in the previous section, the connected operators act by preserving or removing connected components. In this section, we present some simple examples of connected operators, namely we describe some attribute filters as introduced by [Breen and Jones, 1996]. These filters are based on an attribute that describes some features of the connected components and that allows one to decide whether a given component is kept or removed in the result of the filter. As connected operators, they satisfy the property of acting on connected components instead of individual points as developed by [Serra, 1982].

Any attribute filter is based on an *attribute function* \mathcal{A} designed to encode some interesting features of any subset of V (*i.e.*, the vertex set of the graph \mathcal{G} which

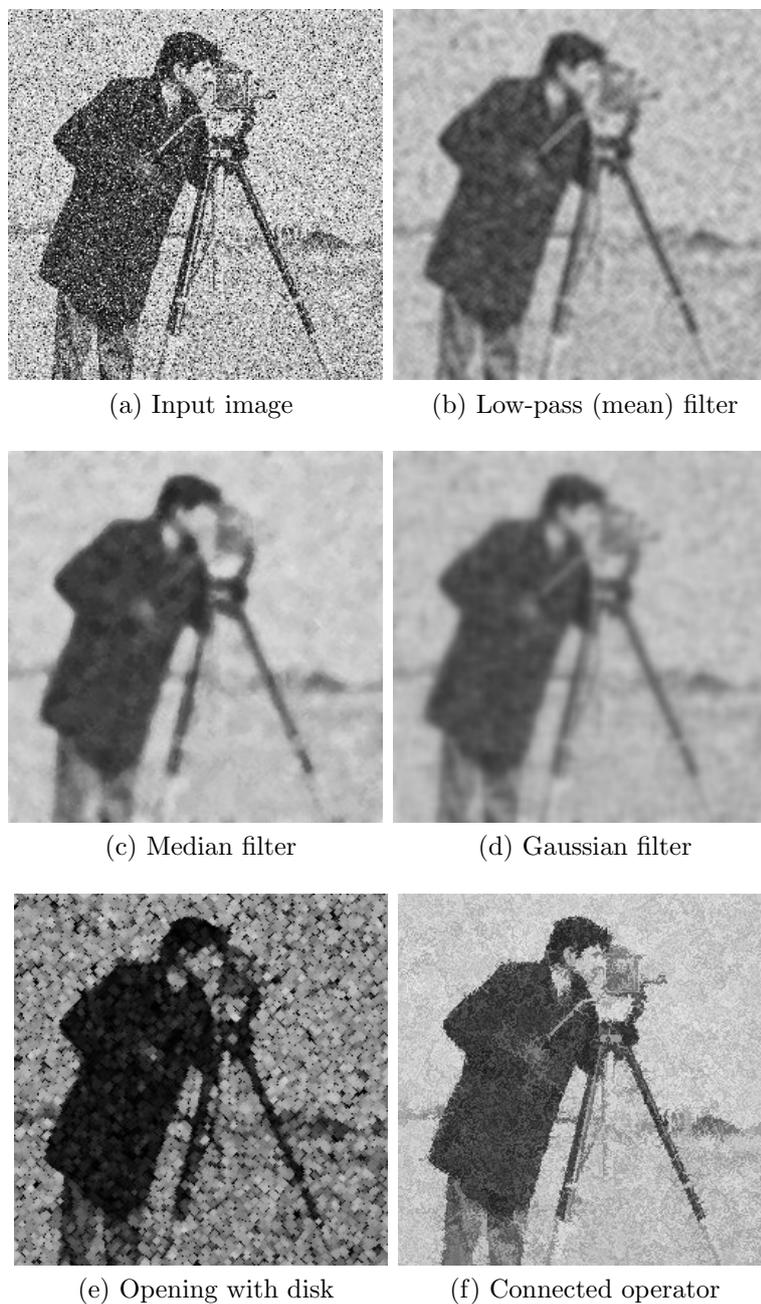


Figure 2.4: A connected operator example on grayscale image for low-level filtering task compared with classical filters: (b) low-pass (mean) filter, (c) median filter, (d) Gaussian filter, (e) opening with structuring element. (f) The grain filter does not create or move contours.

corresponds to the image domain). Such attribute function \mathcal{A} , which is then a mapping from the set \mathcal{V} of all subsets of vertices of \mathcal{G} in the set \mathbb{R}^+ or \mathbb{N} . The attribute function \mathcal{A} can be some shape information measurement such as the area, elongation, perimeter, circularity, compactness, eccentricity or some various notions of shape-complexity [Salembier et al., 1998; Urbach et al., 2007]. In the latter case, the attributes are called shape attributes since they do not depend on the gray levels of the points inside the regions. The family of shape filters [Urbach et al., 2007], *i.e.* the attribute filters obtained from shape-attributes. It has to be noted that the operators called shape-filters in the literature are not always filters in the sense of Section 2.2.3.1 since the increasing property is sometimes relaxed. For this reason some authors such as [Breen and Jones, 1996] prefer to call such operators as attribute thinning. Some attribute functions based on the gray levels of points might also be interesting. It can be as simple as the minimum/maximum/average gray level of the pixels in the considered region or it can be such more complex measure such as the height or volume of the region. As a sample of such attribute, we formally define below three of the most used attribute function: the area, height and volume attributes denoted by $\mathcal{A}_{\text{area}}$, $\mathcal{A}_{\text{height}}$, and, $\mathcal{A}_{\text{volume}}$, respectively. For a given subset of points $C \in \mathcal{V}$, they are given by:

$$\begin{aligned} \mathcal{A}_{\text{area}}(C) &= |C| \\ \mathcal{A}_{\text{height}}(C) &= \max_{x \in C} f(x) - \min_{x \in C} f(x) \\ \mathcal{A}_{\text{volume}}(C) &= \sum_{x \in C} |\max_{x \in C} f(x) - f(x)| \end{aligned} \quad (2.3)$$

Given an attribute function \mathcal{A} and a scalar value λ in \mathbb{R}^+ , we define the *binary connected filter of size λ by the attribute \mathcal{A}* , denoted by $\Gamma_{\mathcal{A}}^{\lambda}$, by

$$\Gamma_{\mathcal{A}}^{\lambda}(X) = \bigcup \left\{ C \in \mathbb{C}(X) \mid \mathcal{A}(C) > \lambda \right\}, \quad (2.4)$$

for any subset X of V .

For instance, we see in Figure 2.5, for a given vertex-weighted graph (\mathcal{G}, f) , the subset X of vertices of \mathcal{G} are shown in red. The filtered subsets for binary connected filters with area, height and volume attributes for different λ is shown in the Figure 2.5b, 2.5c, 2.5d respectively.

The above mentioned attribute filters are indeed filters. When the considered attribute is not increasing, the operator is not necessarily a filter: it is then a binary attribute thinning. A typical example of a binary attribute thinning is the operator obtained from an ovalness attribute, which is not increasing. An illustration of this operator is provided on Figure 2.1. We can easily see that the provided examples of attribute filters are anti-extensive operators. Extensive attribute filters for binary images can be obtained by duality from the anti-extensive versions defined above ($X \rightarrow [\Gamma_{\mathcal{A}}^{\lambda}(X^c)]^c$).

The operators investigated in this section are binary increasing operators. As such, they all induce stack operators acting on functions weighting the vertices of a graph (hence acting on grayscale image). The interested reader can for in-

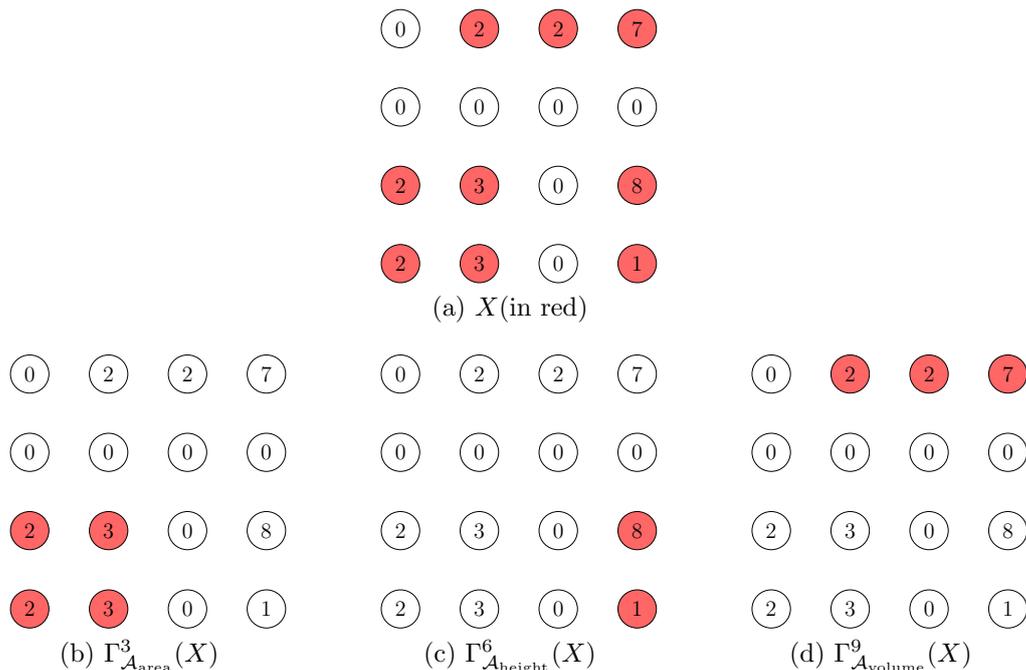


Figure 2.5: Attribute filters

stance refer to [Wendt et al., 1986] for stack operators, to [Serra, 1982; Maragos and Schafer, 1987; Heijmans, 1991; Ronse, 2006] for stack operators in the context of flat mathematical morphology, and to [Bertrand, 2005] for stack operators in the context of watershed image segmentation. This allows filters for weighted graphs and thus for grayscale images to be systematically inferred from the ones on non-weighted graphs. An fruitful presentation of grayscale connected operators obtained by stacking is obtained through the notions of hierarchies and component trees as introduced in the next sections. This structure also allows to go beyond stack operators and to define operators which cannot be obtained by stacking from binary connected operators.

2.2.4 Connectivity in weighted graphs

In this section, we introduce the necessary materials to handle weighted graphs with the particular goal of presenting stack operators obtained from binary connected filters. In Section 2.2.4.1 we recall definitions for edge- and vertex-weighted graphs. Then, we introduce the fundamental structure of a stack of graphs, which allows to handle within a same setting both stack operators for edge- and vertex-weighted graphs. Finally, we present how stacks of graphs can be induced by weighted graphs.

2.2.4.1 Weighted graphs

In graph theory, the traditional weighted graphs consist of weights on edges only. Although, edges have many practical applications, weighing vertices can also serve many purposes. In particular, in image processing a digital image is often seen as

a vertex-weighted graph.

Definition 2.2.7 (Vertex-/Edge-weighted graph). Let $\mathcal{G} = (V, E)$ be a graph, if f (resp. w) is a map from the vertex (resp. edge) set of \mathcal{G} to set \mathbb{R}^+ of positive real numbers, then the pair (\mathcal{G}, f) (resp. (\mathcal{G}, w)) is said to be *vertex-(resp. edge-) weighted graph*.

Any vertex-(resp. edge-) weighted graph can be easily decomposed into a series of sets (of vertices or edges) by thresholding, allowing the elements of the series to be handle with the binary operators presented previously. Let us start with formal definitions of such thresholding operations.

Let $f : X \rightarrow \mathbb{R}^+$ be a function assigning a real positive value to each element of a set X . Given any real positive value $h \in \mathbb{R}^+$, the *lower-level set (or lower threshold) of f at level h* is the set $\mathcal{X}_h(f)$ of elements of X whose value for f is below h :

$$\mathcal{X}_h(f) = \{x \in X \mid f(x) \leq h\} \quad (2.5)$$

Dually, the *upper-level set (or upper threshold) of f at level h* is the set $\mathcal{X}^h(f)$ of elements of X whose value for f is above h :

$$\mathcal{X}^h(f) = \{x \in X \mid f(x) \geq h\} . \quad (2.6)$$

Important notation. In the remaining part of this chapter, the symbol f denotes a mapping from the vertex set V of \mathcal{G} to the set \mathbb{R}^+ of real positive values and the map w denotes a mapping from the edge set E of \mathcal{G} to the set \mathbb{R}^+ . Therefore, the pairs (\mathcal{G}, f) and (\mathcal{G}, w) are vertex- and edge-weighted graphs respectively.

2.2.4.2 Stacks of graphs

In order to understand the notion of connectivity in weighted graphs, we present the stacks of graphs [Perret et al., 2015]. We will then see several ways to induce stacks of graphs from weighted graphs, leading to several notions of connected components of a weighted graphs allowing the design of connected operators for weighted graphs.

Definition 2.2.8 (Stack of graphs). A stack of graphs is a finite sequence $\mathcal{S} = (\mathcal{G}_0, \dots, \mathcal{G}_\ell)$ of graphs such that, for any i in $\{1, \dots, \ell\}$, the graph \mathcal{G}_i is a subgraph of \mathcal{G}_{i-1} . We say that \mathcal{G}_i is the *level set of \mathcal{S} at level i* , for any $i \in \{1, \dots, \ell\}$.

Definition 2.2.9 (Component of a stack). Let $\mathcal{S} = (\mathcal{G}_0, \dots, \mathcal{G}_\ell)$ be a stack. A (*connected*) *component of \mathcal{S}* is a pair (i, X) such that X is connected component of the level set of \mathcal{S} at level i . The stack \mathcal{S} is *connected* when \mathcal{S}_0 is connected.

An illustration of stack of graph can be seen in Figure 2.6, which shows a connected stack composed of four graphs $(\mathcal{G}_0, \dots, \mathcal{G}_3)$.

As seen in previous section, a binary connected operators act by removing connected components from a set of vertices. This notion can be easily extended to stacks. Intuitively, a connected operators on a stack can be thought of as an operator which removes certain connected components of the stack, for instance based

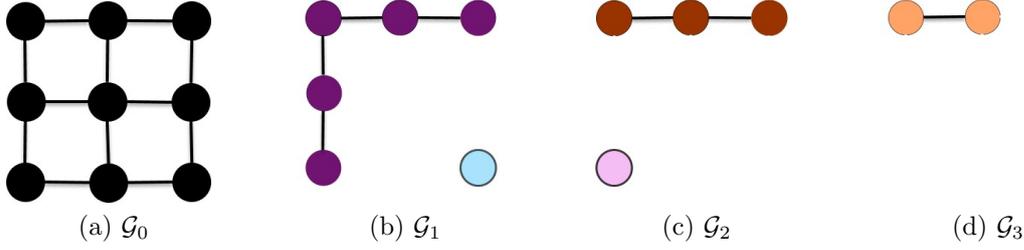


Figure 2.6: A stack $\mathcal{S} = (\mathcal{G}_0, \mathcal{G}_1, \mathcal{G}_2, \mathcal{G}_3)$. Each color represent a component of the stack \mathcal{S} .

on a shape attribute on the components. Before studying such filters with more details, let us describe some means to induce a stack from a weighted graph (Section 2.2.4.3 and 2.2.4.4) and study some hierarchical properties of the set of all connected components of a stack (Section 2.2.5).

A graph may be weighted either on vertices or on edges. Classically, when dealing with connected operators, these two cases are handled separately. However, in this work, based on the notion of a stack, we tackled both in a same setting. This allows us to highlight the common structures of operators which otherwise appear weakly- or even un-related. ¹

2.2.4.3 Stacks induced by a vertex-weighted graph

Intuitively, in order to obtain a stack from the vertex-weighted graph (\mathcal{G}, f) , we consider the decomposition of f into its upper-level or lower-level sets and we consider the graphs induced by these levels sets to form a stack. We recall that, if P is a subset of vertices of \mathcal{G} , then $\epsilon(P)$ contains every edge of \mathcal{G} whose vertices are in P .

Definition 2.2.10 (Vertex induced stacks). The *stack lower-induced by f on V* and the *stack upper-induced by f on V* , denoted by $\mathcal{S}_{\bullet}^{\leq}(f)$ and $\mathcal{S}_{\bullet}^{\geq}(f)$ respectively, are defined by:

$$\mathcal{S}_{\bullet}^{\leq}(f) = \left(\left(\mathcal{X}_h(f), \epsilon(\mathcal{X}_h(f)) \right) \mid h \in \mathbb{R}_f^{\geq} \right), \text{ and} \quad (2.7)$$

$$\mathcal{S}_{\bullet}^{\geq}(f) = \left(\left(\mathcal{X}^h(f), \epsilon(\mathcal{X}^h(f)) \right) \mid h \in \mathbb{R}_f^{\leq} \right), \quad (2.8)$$

where \mathbb{R}_f^{\geq} (resp. \mathbb{R}_f^{\leq}) is the series of values of f ranked in decreasing order (resp. increasing order).

An illustration of stack induced by vertex-weighted graph is shown in the Figure 2.7. It can be observed that both the stacks $\mathcal{S}_{\bullet}^{\leq}(f)$ and $\mathcal{S}_{\bullet}^{\geq}(f)$ are connected.

It can be observed that the *stack lower-induced by f on V* is indeed a stack since $(\mathcal{X}_h(f), \epsilon(\mathcal{X}_h(f)))$ is a subgraph of $(\mathcal{X}_{h'}(f), \epsilon(\mathcal{X}_{h'}(f)))$ for any $h' > h$. Similarly, the

¹Note also that the framework of stacks allows one to deal with non-symmetric graphs and to build hierarchical image representations where asymmetric information can be taken into account and to obtain connected filters relying on this information [Perret et al., 2015].

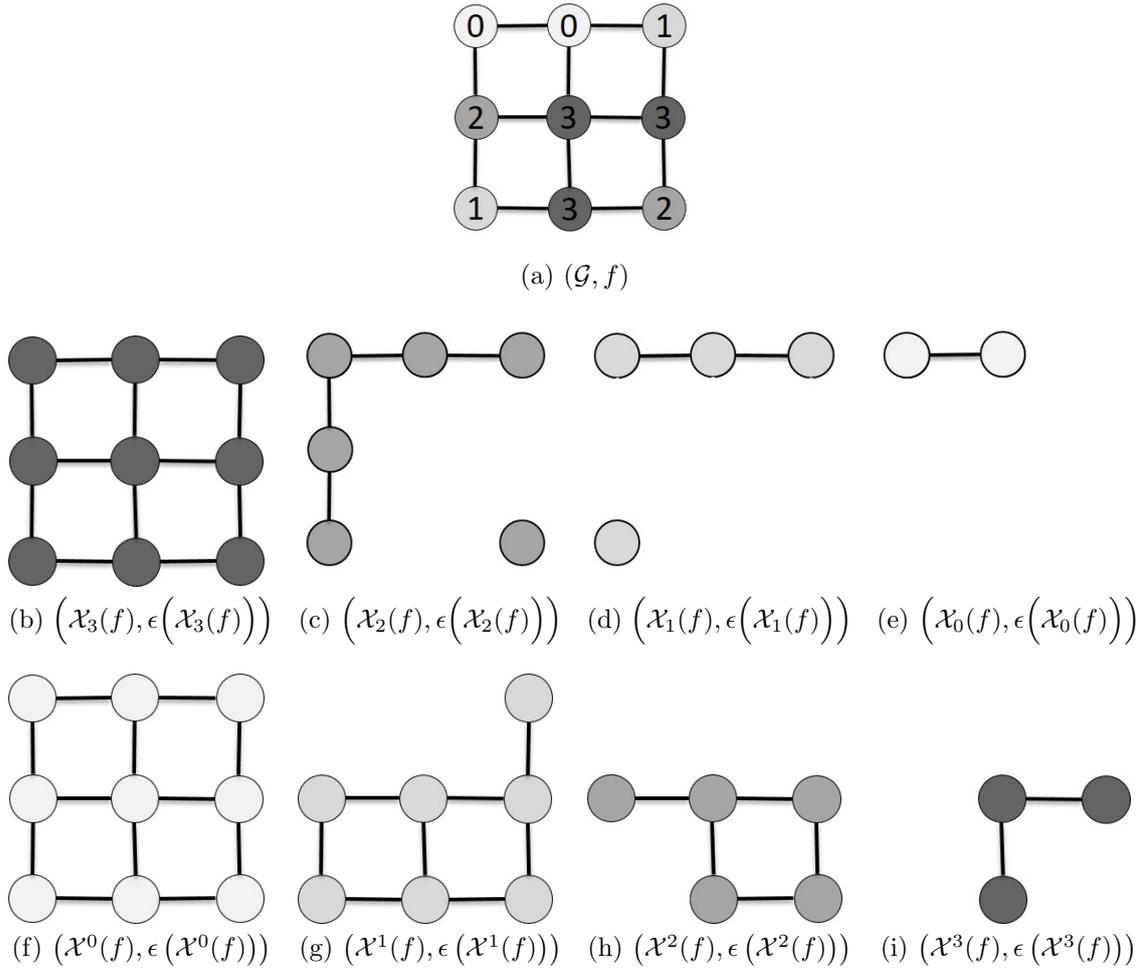


Figure 2.7: An illustration of stack lower-induced and upper-induced by a vertex weighted graph shown in (a). The stack lower-induced $\mathcal{S}_{\bullet}^{\leq}(f)$ is the series of graphs ((b), (c), (d), (e)). The stack upper-induced $\mathcal{S}_{\bullet}^{\geq}(f)$ is the series of graphs ((f), (g), (h), (i)).

stack upper-induced by f on V is also a stack since $(\mathcal{X}^h(f), \epsilon(\mathcal{X}^h(f)))$ is a subgraph of $(\mathcal{X}^{h'}(f), \epsilon(\mathcal{X}^{h'}(f)))$ for any $h' < h$.

2.2.4.4 Stacks induced by an edge-weighted graph

In order to obtain a stack from an edge-weighted graph, we provide a similar construction, namely we consider the series of all level sets and associate a graph with each of these level sets. Note that any level set $X = \mathcal{X}_h(w)$ of (\mathcal{G}, w) is a subset of the edge set of \mathcal{G} . In such case, two interesting subgraphs of \mathcal{G} can be considered.

1. The first one is the graph $(\delta(X), X)$ whose vertex set contains every vertex of \mathcal{G} which belongs to an edge in X and whose edge set is X itself.
2. The second one is the graph (V, X) whose vertex set contains all vertices of \mathcal{G} and whose edge set is X .

Considering also that we can consider both upper and lower level sets, we distinguish four possibilities for a stack induced by an edge-weighted graph.

Definition 2.2.11 (Edge induced stacks). The *stack lower-induced*, the *stack complete-lower-induced*, the *stack upper-induced*, and the *stack complete-upper-induced* by w on V , denoted by $\mathcal{S}_{\times}^{\leq}(w)$, $\mathcal{S}_{\times}^{\leq}(w)$, $\mathcal{S}_{\times}^{\geq}(w)$, $\mathcal{S}_{\times}^{\geq}(w)$, are defined by:

$$\mathcal{S}_{\times}^{\leq}(w) = \left((\delta(\mathcal{X}_h(w)), \mathcal{X}_h(w)) \mid h \in \mathbb{R}_w^{\geq} \right), \quad (2.9)$$

$$\mathcal{S}_{\times}^{\leq}(w) = \left((V, \mathcal{X}_h(w)) \mid h \in \mathbb{R}_w^{\geq} \right), \quad (2.10)$$

$$\mathcal{S}_{\times}^{\geq}(w) = \left((\delta(\mathcal{X}^h(w)), \mathcal{X}^h(w)) \mid h \in \mathbb{R}_w^{\leq} \right), \text{ and} \quad (2.11)$$

$$\mathcal{S}_{\times}^{\geq}(w) = \left((V, \mathcal{X}^h(w)) \mid h \in \mathbb{R}_w^{\leq} \right), \quad (2.12)$$

where \mathbb{R}_w^{\geq} (resp. \mathbb{R}_w^{\leq}) is the series of values of f ranked in decreasing order (resp. increasing order).

An illustration of these four types of stacks induced by edge-weighted graph is shown in Figure 2.8. Similar to stacks induced by vertex-weighted graph, the stacks $\mathcal{S}_{\times}^{\leq}(w)$, $\mathcal{S}_{\times}^{\leq}(w)$, $\mathcal{S}_{\times}^{\geq}(w)$ and $\mathcal{S}_{\times}^{\geq}(w)$ are connected.

It can be noted that the stack lower-induced, the stack complete-lower-induced, the stack upper-induced, and the stack complete-upper-induced by w on V are all indeed stacks.

The notations \bullet and \times in different stack of graphs $\mathcal{S}_{\bullet}^{\geq}(f)$, $\mathcal{S}_{\times}^{\geq}(w)$, $\mathcal{S}_{\bullet}^{\leq}(f)$, $\mathcal{S}_{\times}^{\leq}(w)$ refers to vertex- and edge-weighted graphs respectively and the notations \leq and \geq refers to lower- and upper-induced stacks respectively. In the next Section (2.2.5) on partitions and hierarchies we will see that these various stack of graphs can be expressed as hierarchies.

2.2.5 Connected partitions and hierarchies

As seen previously connected operators help bridging the gap between image segmentation and filtering. So far, we have seen connected operators from the filtering point of view. In this section, we introduce the (hierarchical) segmentation counterpart. Indeed, many image segmentation methods look for a partition of the set of image pixels such that each region of the partition corresponds to an object of interest in the image. Although finding a single partition of an image is still an active topic, it is now recognized that a more robust approach is working in a multi-scale approach that can be given in the form of a hierarchy [Soille, 2008]. Hierarchical segmentation methods, instead of providing a unique partition, produce a sequence of nested partitions at different scales, allowing the description of an object of interest as a grouping of several objects of interest that appear at lower scales. The stacks of graphs introduced in the previous sections allows such interesting hierarchical segmentation to be obtained.

In this section, we define the basic notions of partitions and hierarchies which will be further used to define hierarchies given by a stack of graphs leading to

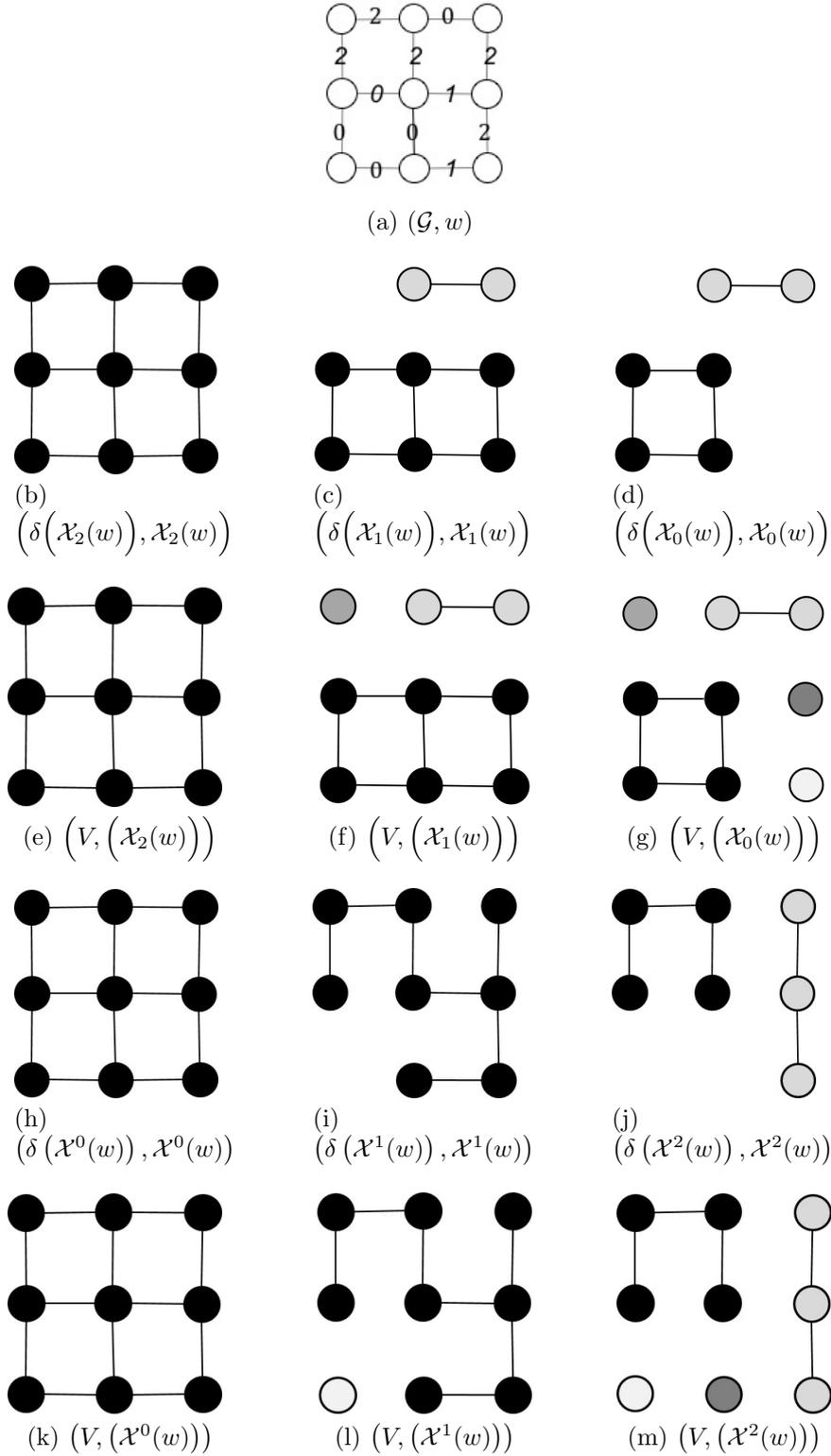


Figure 2.8: An illustration of stack lower-induced, the stack complete-lower-induced, the stack upper-induced, and the stack complete-upper-induced by edge-weighted graph shown in Figure (a). The stack lower-induced $\mathcal{S}_{\times}^{\leq}(w)$ is the series of graphs ((b), (c), (d)), the stack complete-lower induced $\mathcal{S}_{\times}^{\leq}(w)$ is the series of graphs ((e), (f), (g)), the stack upper-induced $\mathcal{S}_{\times}^{\geq}(w)$ is the series of graphs ((h), (i), (j)) and the stack complete-upper induced is the series of graphs ((k), (l), (m)).

the well known notions of component trees which are tree-based representations of images.

Definition 2.2.12 (Partition). A *partition of V* is a set \mathbf{P} of non-empty mutually disjoint subsets of V whose union is V , that is

- for any two X and Y in \mathbf{P} such that $X \neq Y$, we have $X \cap Y = \emptyset$; and
- $\cup\{X \in \mathbf{P}\} = V$.

Definition 2.2.13 (Partial partition). A *partial partition of a V* is a partition \mathbf{P}^* of any subset V_* of V , that is

- for any two X and Y in \mathbf{P} such that $X \neq Y$, we have $X \cap Y = \emptyset$; and
- $\cup\{X \in \mathbf{P}\} \subseteq V$.

Any element X of a (partial) partition \mathbf{P} of V is called a *region of \mathbf{P}* . It can be seen that any partition is a partial partition and that for any partial partition \mathbf{P} , there is unique region of \mathbf{P} containing $x \in V$.

Let \mathcal{G}' be a graph. The *connected component partition of \mathcal{G}'* , denoted by $\mathbb{C}(\mathcal{G}')$ is the set which contains the vertex set of every connected component of \mathcal{G}' :

$$\mathbb{C}(\mathcal{G}') = \{V(\mathcal{C}) \mid \mathcal{C} \text{ is a connected component of } \mathcal{G}'\}.$$

It is well known that $\mathbb{C}(\mathcal{G}')$ is a partition of $V(\mathcal{G}')$. Thus, if \mathcal{G}' is a subgraph of \mathcal{G} , then $\mathbb{C}(\mathcal{G}')$ is a partial partition of V . Furthermore, whenever we have $V(\mathcal{G}') = V$, the set $\mathbb{C}(\mathcal{G}')$ is a partition of V .

We can consider a partial order relationship between (partial) partitions: given two (partial) partitions \mathbf{P}_1 and \mathbf{P}_2 of set V , partition \mathbf{P}_1 is said to be a *refinement of \mathbf{P}_2* (written as $\mathbf{P}_1 \sqsubseteq \mathbf{P}_2$) if any region of \mathbf{P}_1 is included in the region of \mathbf{P}_2 (i.e. $\forall X \in \mathbf{P}_1, \exists Y \in \mathbf{P}_2, X \subseteq Y$). The union of all regions of a partial partition \mathbf{P}^* is called the *support of \mathbf{P}^** and is denoted by $\text{supp}(\mathbf{P}^*)$; thus \mathbf{P}^* is partition of $\text{supp}(\mathbf{P}^*)$. Thereby, a partition of V is a partial partition whose support is V .

Definition 2.2.14 (Hierarchy). A hierarchy of partitions (resp. partial partitions) on V is a sequence $\mathcal{H} = (\mathbf{P}_0, \dots, \mathbf{P}_\ell)$ of partitions of V (resp. partial partitions of V), such that \mathbf{P}_{i-1} is a refinement of \mathbf{P}_i , for any $i \in \{1, \dots, \ell\}$.

If $\mathcal{H} = (\mathbf{P}_0, \dots, \mathbf{P}_\ell)$ is a hierarchy, the integer ℓ is called the *depth* of the hierarchy \mathcal{H} . A *hierarchy $\mathcal{H} = (\mathbf{P}_0, \dots, \mathbf{P}_\ell)$ is called complete* if $\mathbf{P}_\ell = \{V\}$ and if \mathbf{P}_0 contains every singleton of V (i.e., $\mathbf{P}_0 = \{\{x\} \mid x \in V\}$). Further details on hierarchies in framework of partial partition are explained by [Ronse, 2014]. A partition of V is *connected* if every of its regions is connected and a hierarchy on V is connected (for \mathcal{G}) if every of its partitions is connected. Further details on connected hierarchies and their representations are explained by [Cousty et al., 2017].

Figure 2.9 represents a hierarchy $\mathcal{H} = (\mathbf{P}_0, \mathbf{P}_1, \mathbf{P}_2, \mathbf{P}_3)$ on a rectangular subset $V \subset \mathbb{Z}^2$ made of 9 dots. Here, we notice that \mathbf{P}_1 is refinement of \mathbf{P}_2 as any

region of \mathbf{P}_1 is included in a region of \mathbf{P}_2 . It can also be seen that hierarchy \mathcal{H} is complete as \mathbf{P}_0 is made of singletons and $\mathbf{P}_3 = \{V\}$ is made of single region that contains all elements. Also here, \mathbf{P}^* represents a partial partition of the rectangular subset V as all the elements of \mathbf{P}^* does not sum up to 9 dots.

Definition 2.2.15 (Component hierarchy of a stack). Let $\mathcal{S} = (\mathcal{G}_0, \dots, \mathcal{G}_\ell)$ be a stack of graphs. The *component hierarchy* of \mathcal{S} is the series of connected components partitions of the level sets of \mathcal{S}

$$\mathcal{H}_{\mathcal{S}} = \left(\mathbb{C}(\mathcal{G}_h) \mid h \in \{0, \dots, \ell\} \right) \quad (2.13)$$

It can be easily seen that $\mathcal{H}_{\mathcal{S}}$ is indeed a hierarchy of partial partitions since, for any h in $\{1, \dots, \ell\}$, the connected component partition $\mathbb{C}(\mathcal{G}_h)$ is a refinement of than $\mathbb{C}(\mathcal{G}_{h-1})$, which can easily established by the fact that the level set \mathcal{G}_h of the stack \mathcal{S} is a subgraph of the level set \mathcal{G}_{h-1} . In general, the component hierarchy of a stack \mathcal{S} is not a hierarchy of partitions of V . For instance, the component hierarchies of the stack $\mathcal{S}_{\bullet}^{\leq}(f)$ of Figure 2.7 is indeed a hierarchy of partial partitions but not a hierarchy of partitions since its level set at level-0, 1 and 2 is not a partition of V . Note that this stack is lower-induced by a vertex-weighted graph. Similar examples can be found for stacks which are upper-induced by a vertex-weighted graph and lower or upper induced by an edge weighted graph. However, any level set of a stack which is complete-lower-induced (or complete-upper-induced) by any edge weighted graph induces a hierarchy of partitions of V . For instance, the connected component hierarchies of the stacks $\mathcal{S}_{\times}^{\leq}(w)$ and $\mathcal{S}_{\times}^{\geq}(w)$ of Figure 2.8 is an example of induced hierarchy of partitions. The following property summarizes these results.

Propety 2.2.1. Let \mathcal{S} be a stack. The following statements hold true:

1. if \mathcal{S} is lower-induced by f , then the component-hierarchy of \mathcal{S} is a hierarchy of partial partitions of V but not, in general, a hierarchy of partitions of V ;
2. if \mathcal{S} is upper-induced by f , then the component-hierarchy of \mathcal{S} is a hierarchy of partial partitions of V but not, in general, a hierarchy of partitions of V ;
3. if \mathcal{S} is lower-induced by w , then the component-hierarchy of \mathcal{S} is a hierarchy of partial partitions of V but not, in general, a hierarchy of partitions of V ;
4. if \mathcal{S} is upper-induced by w , then the component-hierarchy of \mathcal{S} is a hierarchy of partial partitions of V but not, in general, a hierarchy of partitions of V ;
5. if \mathcal{S} is complete-lower-induced by w , then the component-hierarchy of \mathcal{S} is a hierarchy of partitions of V ;
6. if \mathcal{S} is complete-upper-induced by w , then the component-hierarchy of \mathcal{S} is a hierarchy of partitions of V .

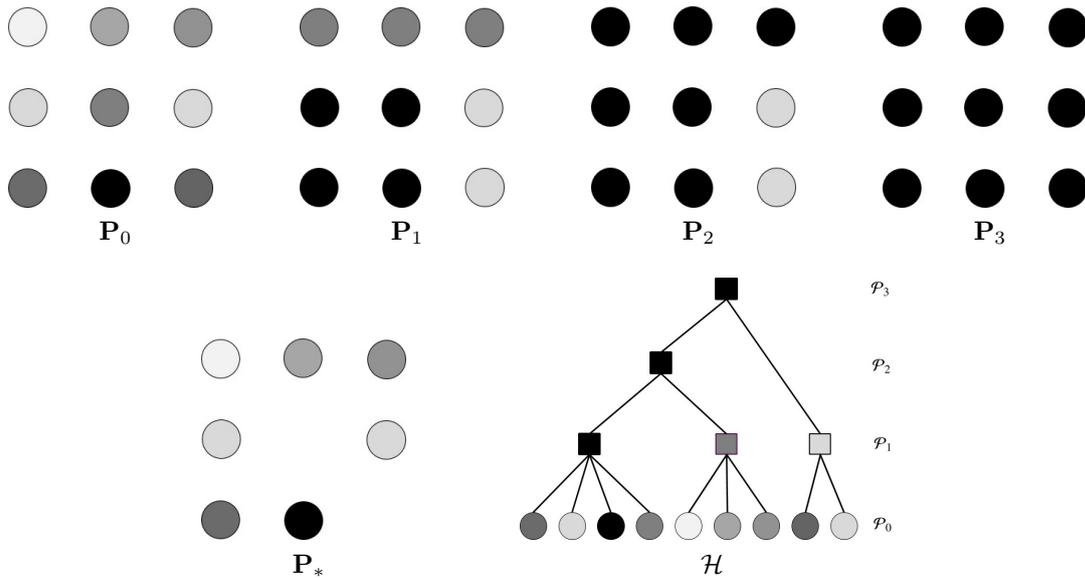


Figure 2.9: An illustration of partition, partial partition and hierarchy on a set V of 9 dots. In first row, for every partition \mathbf{P}_i , each region is represented by a gray level: two dots with same gray level belong to same region. \mathbf{P}_* is the partial partition of the set V . The right subfigure in second row represents the hierarchy $\mathcal{H} = (\mathbf{P}_0, \mathbf{P}_1, \mathbf{P}_2, \mathbf{P}_3)$ as a tree structure, called as a dendrogram, where inclusion relation between regions of successive partitions is represented by line segments.

The component hierarchies of these six stacks are well studied in mathematical morphology, through the notions of min/max trees of an edge and vertex-weighted graphs [Salembier et al., 1998; Najman and Couprie, 2006; Carlinet and Géraud, 2015] and through the notion of quasi-flat zones of an edge-weighted graph [Meyer and Maragos, 2000; Cousty, Najman and Perret, 2013; Bosilj et al., 2018]. In the next section, we present a tree data structure, called a component tree or a tree-based shape space, that can be associated to any of these hierarchies, which allows one to efficiently handle them in computerized procedures for segmenting and filtering images.

2.2.6 Component trees: tree based shape space

A component tree is a compact representation of a hierarchy which is adopted for handling efficiently a hierarchy with computerized procedure. Tree based representations of data have been a classical approach across different fields of mathematics. In discrete mathematics it has been Hasse diagram of partial ordering relations, in classification as the dendrograms of hierarchy and component trees in mathematical morphology.

Let $\mathcal{H} = \{\mathbf{P}_0, \dots, \mathbf{P}_\ell\}$ be a hierarchy. We say that a subset R of V is a *region* of \mathcal{H} if $\exists i \in \{0, \dots, \ell\}$ such that R belongs to \mathbf{P}_i . We denote by $R_{\mathcal{H}}$, the set of all regions of \mathcal{H} . Let $\mathcal{S} = (\mathcal{G}_0, \dots, \mathcal{G}_\eta)$ be a stack of graphs. When the considered hierarchy \mathcal{H} is the component hierarchy $\mathcal{H}_{\mathcal{S}}$ of a stack \mathcal{S} (refer Definition 2.2.15),

any region R of \mathcal{H}_S is an element of the connected component partition of the graph \mathcal{G}_h of the stack \mathcal{S} , *i.e.* $R \in \mathbb{C}(\mathcal{G}_h)$ such that $h \in \{0, \dots, \eta\}$.

The component tree representation is composed of regions (*i.e.* connected components) from small to large, where two regions are either nested or disjoint. In component tree representation, any region is the *node of the component tree*. Let R_1 and R_2 be any two distinct regions, we have a nesting property:

$$R_1 \cap R_2 \neq \emptyset \rightarrow R_1 \subset R_2 \text{ or } R_2 \subset R_1 \quad (2.14)$$

As any tree, the component tree of \mathcal{H} can be defined as a pair made of a set of nodes and of a binary (parent) relation on the set of nodes. More precisely, the *component tree of \mathcal{H}* is the pair $\mathcal{T} = (R_{\mathcal{H}}, \text{parent})$ such that $R_{\mathcal{H}}$ is the set of all regions of \mathcal{H} and such that a region R_1 in $R_{\mathcal{H}}$ is a parent of a region R_2 in $R_{\mathcal{H}}$ whenever R_1 is a minimal (for inclusion relation) proper superset of R_2 . Note that every region in $R_{\mathcal{H}}$ has exactly one parent except the region V which has no parent and is called the *root of the component tree* of \mathcal{H} . Any region which is not the parent of another one is called a *leaf of the tree*. It can be observed that any singleton of V is a leaf of \mathcal{T} and that conversely any leaf of \mathcal{T} is a singleton of V .

Remark: In case of hierarchical representation of images, a tree-based shape spaces $S_{\mathcal{T}}$ [Xu, 2013] is defined as set $\{R_i\}$ of regions, that can be organized into a tree structure \mathcal{T} for which the root node represents R_0 , the entire image domain and any pair of regions R_1 and R_2 follows the nesting property (Equation. 2.14). The neighborhood of the shape space is defined by the inclusion relationship between the two regions. A tree-based shape space is thus a component tree, when the considered graph is image.

A tree-based shape space $S_{\mathcal{T}}$ is built from a tree-based image representation, which is equivalent to image in the sense that the image can be reconstructed from the set of the nodes of the tree. Therefore, tree-based shape space is equivalent tree representation.

A tree-based shape space $S_{\mathcal{T}}$ can also be considered and processed as a connected graph, where the nodes of the tree are vertices of the graph weighted by some attribute function \mathcal{A} and the edges of the graph are the links created by the inclusion relationship between the nodes in the shape space. One of the significant step involves defining the attribute function \mathcal{A} according to the input image and application. Attribute filters for binary images have been described in the Section 2.2.3.3. In case of grayscale images, the attribute function based image filtering, when the tree-representation is the component tree, is the extension of attribute filters to grayscale images.

The different hierarchies induced by stack of graphs mentioned in the previous section forms various component trees which are studied vastly in literature. We first present Min/Max tree which is one of the simplest tree based image representation.

2.2.6.1 Min/Max tree.

The max-tree was firstly proposed by [Salembier et al., 1998] as a suitable and efficient structure to deal with the processing steps involved in antiextensive connected

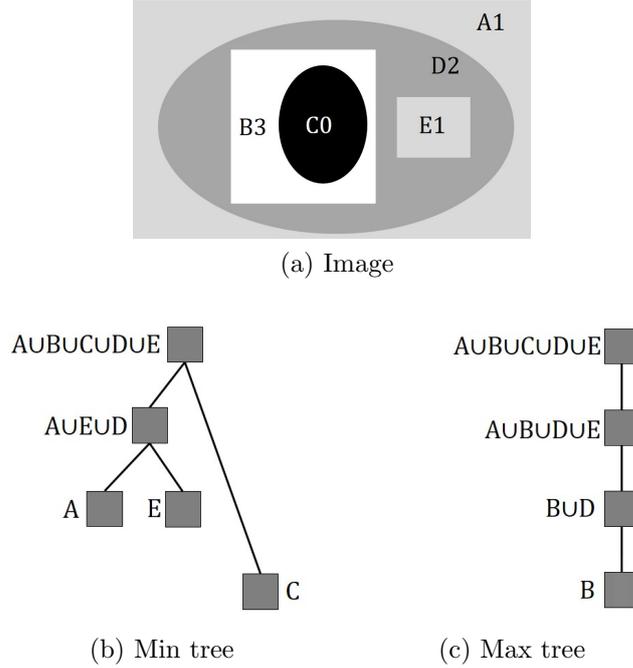


Figure 2.10: An illustration of component tree on an image. (a) Original image where the numbers represent the grayscale value of flat zone. (b) and (c) represent the min tree and max tree respectively where the root is at the top.

operators. The resulting component tree when the considered hierarchy is component hierarchy of the stack $\mathcal{S}_{\bullet}^{\leq}(f)$ (resp. $\mathcal{S}_{\times}^{\leq}(w)$) lower-induced by a vertex- (resp. edge-) weighted graph is the *min tree* of the vertex- (resp. edge-) weighted graph and is denoted by $\mathcal{T}_{\bullet}^{\leq}$ (resp. $\mathcal{T}_{\times}^{\leq}$). Similarly, the component tree when the considered hierarchy is component hierarchy of the stack $\mathcal{S}_{\bullet}^{\geq}(f)$ (resp. $\mathcal{S}_{\times}^{\geq}(w)$) upper-induced by a vertex- (resp. edge-) weighted graph is the *max tree* of the vertex- (resp. edge-) weighted graph and is denoted by $\mathcal{T}_{\bullet}^{\geq}$ (resp. $\mathcal{T}_{\times}^{\geq}$).

In most of the image processing applications, grayscale image f is considered as a vertex-weighted graph (\mathcal{G}, f) , where f is mapping from vertex set V to \mathbb{R}^+ . The min tree of the graph (\mathcal{G}, f) is called the *min tree* $\mathcal{T}_{\bullet}^{\leq}$ of the image f . By duality, the max tree of (\mathcal{G}, f) obtained by upper-level set decomposition leads *max tree* $\mathcal{T}_{\bullet}^{\geq}$ of the image f . By construction, in case of min tree of image, the leaf nodes of the tree are the dark connected components whereas in case of max tree of the image the leaf nodes are the bright connected components. In both these cases, the root node represents the complete image domain V . These min/max tree of images are studied vastly in literature [Salembier et al., 1998; Jones, 1999; Jalba et al., 2006; Westenberg et al., 2007; Berger et al., 2007]. However, the min tree $\mathcal{T}_{\times}^{\leq}$ and max tree $\mathcal{T}_{\times}^{\geq}$ of edge-weighted graphs are not studied in literature.

A simple example of min/max tree is illustrated in Figure 2.10. The original image is composed of five flat zones identified by letters A, B, C, D and E. The numbers next to the letters represent the grayscale value of the flat zones. In our example the grayscale values range from zero to three. The leaves of the min tree corresponds to the regional minima of the image. In this example, as the image has

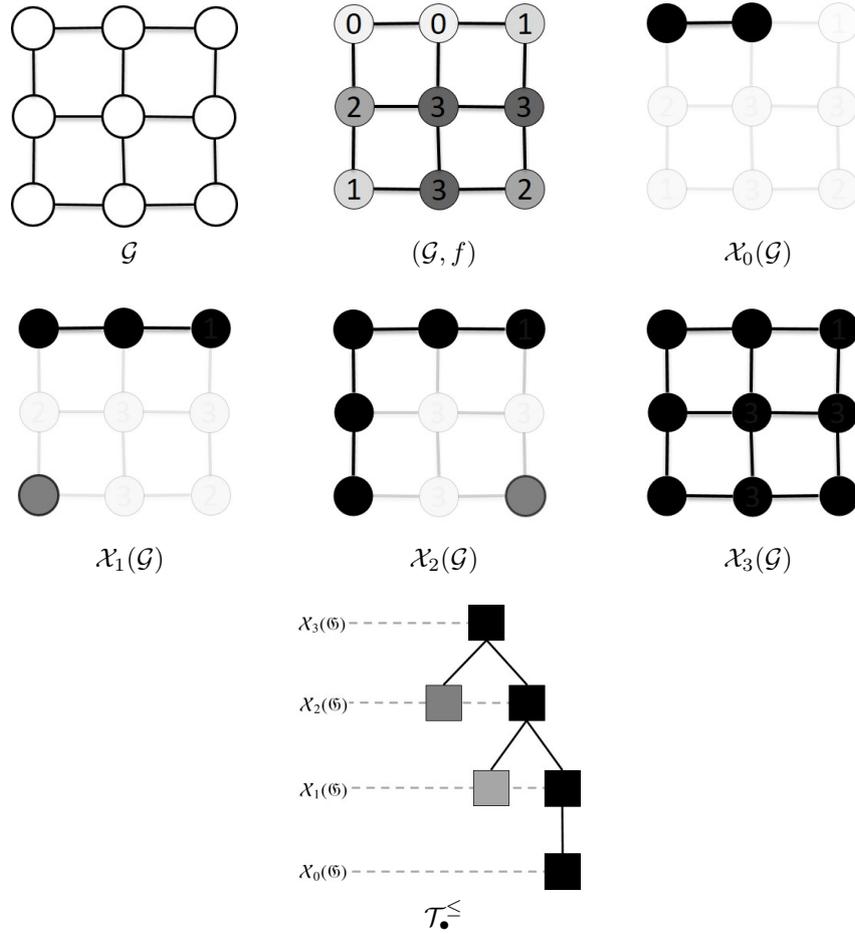


Figure 2.11: An illustration of min tree as a hierarchy of partial partition. (a) a graph; (b) its vertex-weighted graph; (c), (d), (e), (f) are the level sets of graph at levels 0, 1, 2 and 3 respectively; (g) is the min tree of graph (\mathcal{G}, f) .

three minima, the min tree has three leaves. The leaves of the max tree are the regional maxima of the image. As the image of Figure 2.10 has only one maxima, the max tree has only one leaf, hence one branch. The image can be reconstructed from the data of the families of upper-level sets (Equation 2.6) and of lower-level sets (Equation 2.5) [Caselles et al., 1999].

$$f(x) = \inf \left\{ h \in \mathbb{R} \mid x \in \mathcal{X}^h \right\} = \sup \left\{ h \in \mathbb{R} \mid x \in \mathcal{X}_h \right\} \quad (2.15)$$

We now present that the min/max trees are hierarchies of partial partition induced by vertex-weighted graph with the help of an example. Let us consider a graph \mathcal{G} as in Figure 2.11a which is weighted on vertices by map $f : V \rightarrow \mathbb{R}^+$ to obtain vertex-weighted graph (\mathcal{G}, f) , as in Figure 2.11b. In this example, f ranges from zero to three. The level set of graph (\mathcal{G}, f) at level 0 yields a subgraph $\mathcal{X}_0(\mathcal{G})$ (Figure 2.11c) which consist of single connected component composed of two vertices of the vertex set V . Therefore, this subgraph is the partition of the subset

of the vertex set V where the connected component is the region of the partition. According to definition 2.2.13, the level set $\mathcal{X}_0(\mathcal{G})$ is a partial partition of $\mathcal{V}(\mathcal{G})$. Figure 2.11d and 2.11e represent the level sets at levels one and two respectively and consist of two connected components (regions) each. It can also be seen that the two connected components of $\mathcal{X}_1(\mathcal{G})$ are refinements of the larger connected component of $\mathcal{X}_2(\mathcal{G})$. Furthermore, partition $\mathcal{X}_2(\mathcal{G})$ is made of single connected component which contains all elements of V . The component hierarchy of the stack $\mathcal{S}_\bullet^\leq(f)$ lower-induced by f forms the min tree \mathcal{T}_\bullet^\leq (Figure 2.11g) and is a hierarchy of the partial partitions.

The min tree of an image has been used extensively in this thesis for segmenting different interventional tools as they appear as dark objects in the fluoroscopic images. The different algorithms to compute the min/max tree are discussed in the Section 2.2.6.3.

2.2.6.2 Quasi-flat zones tree

The min/max tree is the hierarchy of the partial partition induced by weighted graph. On the contrary, the component tree called as quasi-flat zones tree [Meyer and Maragos, 2000] is the hierarchy of the partition of edge-weighted graph. Quasi-flat zones tree has been widely used in image processing and is also known as hierarchy of single linkage [Nagao et al., 1979], or constrained connectivity [Soille, 2008], and α -tree [Ouzounis and Soille, 2011].

As seen in Section 2.2.4.4, for an edge-weighted graph (\mathcal{G}, w) , the stack complete lower- (upper-) induced $\mathcal{S}_\times^\leq(w)$ (resp. $\mathcal{S}_\times^\geq(w)$) on E is series of graphs whose vertex set is same as that of (\mathcal{G}, w) . Hence, unlike the min/max tree, the graph induced by each level-set $\mathcal{X}_h(w)$ is the partition of V . The corresponding hierarchy \mathcal{H}_\times^\leq (or \mathcal{H}_\times^\geq) is thus the hierarchy of the partition induced by the level-sets of edge-weighted graph. The resulting component tree when the considered hierarchy is component hierarchy of the stack $\mathcal{S}_\times^\leq(w)$ (or $\mathcal{S}_\times^\geq(w)$) lower- (or upper-) induced by an edge-weighted graph is the *quasi-flat zones tree* of the edge-weighted graph and is denoted by \mathcal{T}_\times .

A simple example of quasi-flat zones tree is given in Figure 2.12. Let us consider the graph \mathcal{G} shown in Figure 3.10 and the map w shown in Figure 2.12b. The 0-, 1-, 2- and 3-level sets of \mathcal{G} contain the edges shown in Figures 2.12c, 2.12d, 2.12e, 2.12f, respectively. The graphs depicted in this figure are the associated 0-, 1-, 2- and 3-level graphs of \mathcal{G} and the associated 0-, 1-, 2- and 3-level partitions are shown in Figure 2.12g.

Two distant maps that weight the edges of the same graph can induce the same quasi-flat zones tree. As we can observe, only the edges present in Figure 2.12f are sufficient to define corresponding hierarchy in Figure 2.12g. Therefore, in this case, some of the edge weights do not convey any useful information with respect to corresponding quasi-flat zones tree. Therefore, for an edge-weighted graph $\mathcal{G}(E)$, its minimum-spanning tree is the minimal subgraph $\mathcal{G}_* \sqsubseteq \mathcal{G}$ such that the quasi-flat zones trees of \mathcal{G} and \mathcal{G}_* are the same. Interested readers are directed to the work of [Cousty et al., 2017] for more details about this link between quasi-flat zones tree

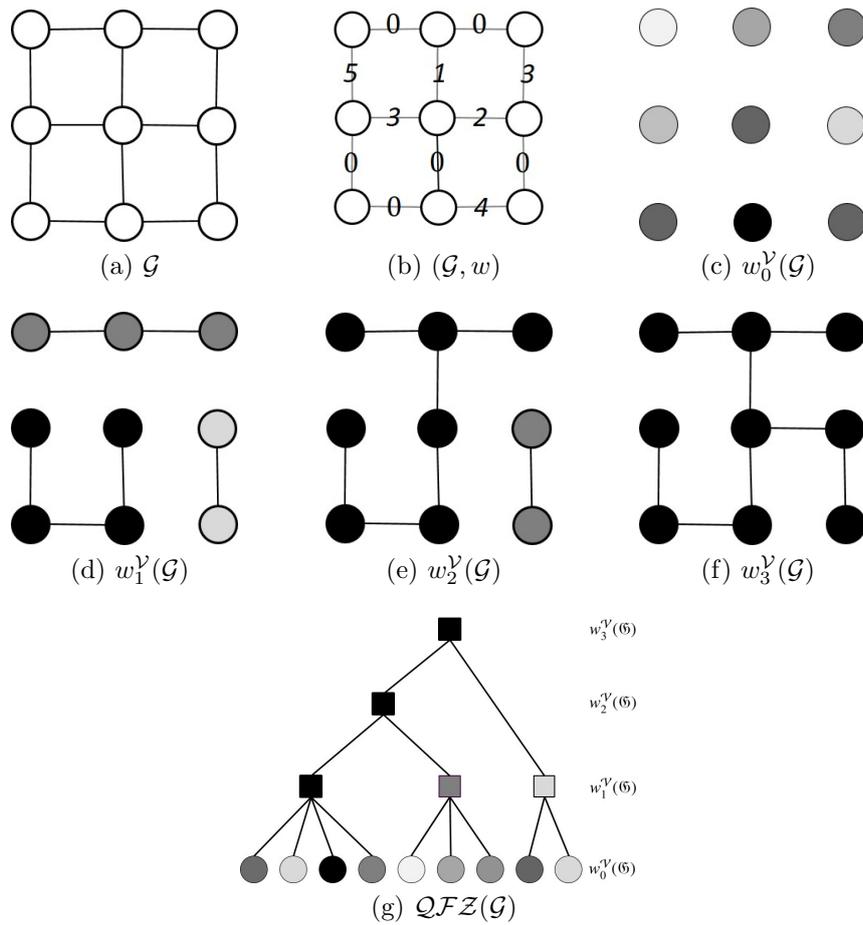


Figure 2.12: Illustration of quasi-flat zones tree. (a) A graph \mathcal{G} ; (b) a edge-weighted graph (\mathcal{G}, w) of \mathcal{G} where edges (in gray) are weighted (numbers in black) with map w ; (c), (d), (e), (f) the λ -level graph of \mathcal{G} , with $\lambda = 0, 1, 2, 3$, respectively; (g) the associated connected component partitions that make up the quasi-flat zones tree of \mathcal{G} for w .

and minimum spanning tree.

For image analysis applications, the edge-weighted graph (\mathcal{G}, w) can be obtained as a pixel or a region adjacency graph: the vertex set of (\mathcal{G}, w) is either the domain of the image to be processed or the set of regions of an initial partition of the image domain. In the latter case, the regions are often called the “image superpixels” (see, e.g., [Achanta et al., 2012]). In both cases, two typical settings for the edge set of \mathcal{G} can be considered:

- the edges of (\mathcal{G}, w) are obtained from an adjacency relation between the image pixels, such as the well known 4- or 8-adjacency relations; and
- the edges of (\mathcal{G}, w) are obtained by considering, for each vertex v of (\mathcal{G}, w) , the nearest neighbors of v for a distance in a features space onto which the vertices of (\mathcal{G}, w) are mapped.

A common feature space (see, e.g., [Felzenszwalb and Huttenlocher, 2004]) is the one where each pixel of a color image is mapped to a vector in dimension 5 made of the two spatial coordinates and the three spectral values describing the color of the pixel. Also, the weight $w(e)$ of an edge between pair of neighboring pixels x and y of image f may be the dissimilarity measure between two pixels. This dissimilarity measure can be as simple as the intensity difference $w(e) = |f(x) - f(y)|$. In some other cases [Guimarães et al., 2017], the edges are weighted by a simple color gradient: the Euclidean distance in the RGB space between the colors of the two adjacent pixels. Some algorithms to compute this tree are mentioned in next section.

2.2.6.3 Algorithms for component tree construction

Several fast algorithms have been proposed in literature to efficiently compute component trees. In this section we briefly discuss the various algorithms used to compute the min/max tree, α -tree (quasi-flat zones trees) and tree of shapes.

The various algorithms used to compute the min/max tree can be classified in to three approaches as compared by [Carlinet and Géraud, 2014]. The first approach is based on flood-filling [Salembier et al., 1998; Nistér and Stewénus, 2008; Wilkinson, 2011] which generally starts at the root and performs a breadth-first or depth-first flooding procedure to make the tree. And the second approach called as immersion algorithms [Najman and Couprie, 2006; Aho et al., 1983; Berger et al., 2007] uses Tarjan’s union-find procedure [Tarjan, 1975], consisting of two passes. A first pass sorts the vertex weights/pixels, and the second pass, build the tree while performing the union-find procedure. This type of algorithm has quasi-linear complexity. The third approach is merge-based algorithms, which divide image in blocks and compute min/max tree on each sub-image using another min/max tree algorithms. Sub min/max trees are then merged to form the tree of the whole image. Those algorithms are well-suited for parallelism using a map-reduce (or divide-and-conquer) approach [Ouzounis and Wilkinson, 2007b; Wilkinson et al., 2008]. When blocks are image lines, dedicated 1D max tree algorithms can be used [Matas et al., 2008; Menotti et al., 2007; Morard et al., 2012].

The α -tree construction algorithm relies on its equivalence with a min tree defined on the edges valued with pixel intensity differences [Soille and Najman, 2012; Najman, 2011], and can use any min tree algorithm. Extending the idea, [Havel et al., 2016, 2013] calculate the α -tree directly using a modification of Tarjan’s union-find [Tarjan, 1975], presenting an algorithm suited for multi-threading applications. The algorithm by [Najman et al., 2013], inspired by Kruskal’s Minimum Spanning Tree algorithm [Kruskal, 1956] and using Tarjan’s union-find [Tarjan, 1975], first constructs a BPT by Altitude Ordering, from which the α -tree can be obtained with a linear post-processing step [Cousty, Najman and Perret, 2013; Najman et al., 2013].

The Tree of Shapes (ToS) [Monasse and Guichard, 2000b,a], also called topographic map, is an inclusion tree combining the min and max tree to represent both bright and dark structures simultaneously. Early approaches to ToS construction operated with worst-case time complexity of $O(N^2)$ [Ballester et al., 2001; Caselles and Monasse, 2009] and were not easily extendible to multidimensional (nD) images [Monasse and Guichard, 2000a]. A recent algorithm by [Géraud et al., 2013] overcomes these drawbacks by using the immersion algorithms for min tree construction as a canvas, and replacing the sorting step.

2.2.7 Filtering on component trees

The reviewed image representations of min/max trees and quasi-flat zones trees can be considered as trees \mathcal{T}_T based on thresholding decompositions and trees \mathcal{T}_H based on hierarchical image segmentation. The major difference between these two representations is that any cut (except the root) on the min/max tree \mathcal{T}_T yields a subset of image domain whereas any cut on the quasi-flat zones tree \mathcal{T}_H gives a partition of the image domain.

A component tree \mathcal{T} is built from a tree-based image representation, which is equivalent to image in the sense that the image can be reconstructed from the set of the nodes of the tree. Therefore, component tree is equivalent tree representation. This component tree \mathcal{T} can also be considered as a connected graph, where the nodes of the tree are vertices of the graph weighted by some attribute function \mathcal{A} and the edges of the graph are the links created by the inclusion relationship between the nodes in the shape space. One of the significant step involves defining the attribute function \mathcal{A} according to the input image and application. Once the tree is constructed and the attribute function \mathcal{A} is defined, the connected operators include a tree filtering step which simplifies it based on the attribute function. Attribute filters for binary images have been described in the Section 2.2.3.3. In case of grayscale images, the attribute function based image filtering, when the tree-representation is the min/max tree, is the extension of attribute filters to grayscale images.

Based on the filtered nodes and attribute functions, the tree filtering strategies can be divided into two categories: pruning strategies and non-pruning strategies. Tree pruning strategies consist of removing of the complete subtree of tree below a considered node for which the considered node is root, while other nodes above

that node are preserved. It can be seen as cutting of sub-branches of the tree. If a node is filtered by pruning strategy, then its descendants are also filtered. For the case of non-pruning strategies, the descendants of the node to be removed may be preserved. We distinguish the attribute function \mathcal{A} based image filtering in the following two sections 2.2.7.1 and 2.2.7.2.

2.2.7.1 Increasing attributes based filtering

An attribute function \mathcal{A} is said to be increasing if,

$$\mathcal{A}(\mathcal{N}) \leq \mathcal{A}(\mathcal{N}_p) \quad \forall \mathcal{N} \in \mathcal{T} \quad (2.16)$$

where \mathcal{N}_p is parent node of node \mathcal{N} according to inclusion relationship. We denote an increasing attribute function by the notation \mathcal{A}^\wedge . Some examples of increasing attributes $\mathcal{A}^\wedge(\mathcal{N})$ for any given node $\mathcal{N} \in \mathcal{T}$ are area, height and volume denoted by $\mathcal{A}_{\text{area}}^\wedge$, $\mathcal{A}_{\text{height}}^\wedge$, and, $\mathcal{A}_{\text{volume}}^\wedge$ (refer Equation 2.3), respectively. Another example of increasing attribute can be the diameter of the smallest (resp. largest) circle that encloses (resp. can fit into) \mathcal{N} [Breen and Jones, 1996].

In case of increasing attributes \mathcal{A}^\wedge , the tree filtering is performed by pruning the nodes whose attribute functions \mathcal{A}^\wedge is less than a given threshold, which can be considered as attribute thresholding. The increasingness of the attribute function make the attribute thresholding a pruning strategy.

2.2.7.2 Non-increasing attributes based filtering

Most of the interesting attribute functions \mathcal{A} , like the shape attributes described in Section 2.2.3.3 which describes the form of the shapes are non-increasing attributes [Salembier et al., 1998; Urbach et al., 2007]. Some examples of non-increasing attribute functions are:

- Perimeter $P(\mathcal{N})$ of the node \mathcal{N} ;
- Compactness(\mathcal{N}) = $\frac{4\pi \text{Area}(\mathcal{N})}{P^2(\mathcal{N})}$, similar attributes computing circularity of the node are presented by [Montero and Bribiesca, 2009];
- Elongation(\mathcal{N}) = $\frac{l_{\max}(\mathcal{N})}{l_{\min}(\mathcal{N})}$, where l_{\max} and l_{\min} are the lengths of the major and minor axes of the best fitting ellipse of the node \mathcal{N} ;
- Sharpness(\mathcal{N}) = $\frac{\text{Volume}(\mathcal{N})}{\text{Area}(\mathcal{N}) \times \text{Height}(\mathcal{N})}$;
- Largest geodesic distance in the connected component represented by \mathcal{N} [Breen and Jones, 1996];
- Geodesic tortuosity and geodesic elongation [Morard et al., 2013].

The attribute thresholding has also been proposed for non-increasing attributes, originally as attribute thinnings by [Breen and Jones, 1996]. In this case a node is removed iff $\mathcal{A}(\mathcal{N}) < t$, where t is the attribute threshold. The content of the removed nodes are merged with their lowest preserved ancestor. Once the node to be removed is selected, the image filtering result depends on either of the two reconstruction rules, namely: 1) direct rule [Salembier et al., 1998], and 2) subtractive rule [Breen and Jones, 1996; Urbach et al., 2007]. Three more pruning strategies (*Min*, *Max*, *Viterbi*) were proposed for non-increasing attributes by [Salembier et al., 1998] and [Urbach et al., 2007].

All these attributes are scalar attributes, *i.e.* the attribute function \mathcal{A} associates a scalar value for each node. More recently, vectorial attributes were proposed, which consist of a vector of scalar attributes. In [Urbach et al., 2005; Perret and Collet, 2015], the authors proposed filtering based on vector attributes where node is removed on the basis of designed distance measure between the attribute vector of the attribute vector of the node and reference vector. There exist many works in literature that perform image segmentation based on “relevant” shape extraction from tree-based shape spaces. For instance, a set of filtering strategies using max tree representation for image segmentation was proposed by [Salembier et al., 1998]. [Najman, 2011] in his work shows that segmentation result can be easily obtained by an horizontal cut on the hierarchy of segmentations. Also, [Felzenszwalb and Huttenlocher, 2004] in their work proposed efficient graph based image segmentation algorithm using the minimum spanning tree. The authors of [Guigues et al., 2006] proposed scale-sets theory which was modeled as an energy minimization problem to obtain non-horizontal optimal cut.

For attribute thresholding based filtering strategies, including the pruning strategy for increasing attributes, the regions of component trees are considered individually based on the attribute value of the nodes themselves. In particular, such simple filtering processes do not take into account the intrinsic parenthood relationship of the tree. Therefore, [Xu, 2013] in his work proposed the framework of shape-based morphology (or shaping), where the idea is to apply connected operators in the tree-based shape space $S_{\mathcal{T}}$. More information on filtering with shaping is detailed in next section 2.2.7.3.

2.2.7.3 Filtering with shaping

The tree-based shape spaces (Definition B.2) allows us to view classical connected operators in a new way. They can be perceived as granular analysis of the shape space without using the tree structure. In [Xu et al., 2016], the authors explain that the idea of *shape-based morphology* or *morphological shaping* is to apply connected filters in the space of all the components of the image, such space being structured into a graph by parenthood relationship (*i.e.*, the neighbors of a node are the parent and the child nodes). This space of all components of image is the tree-based shape space. This framework consists of two trees construction: the first one is constructed from the image, and the second one is constructed from the first tree representation. This process is explained by black+red path in Figure 2.13. The process depicted in the red block is exactly a tree-based connected operator on the tree-based shape

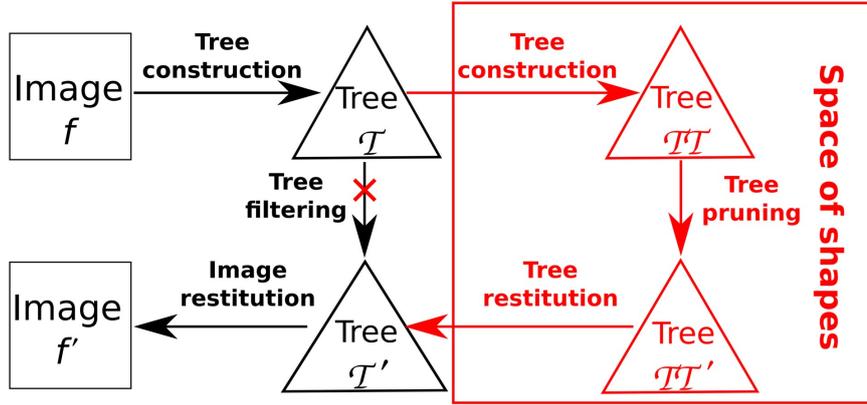


Figure 2.13: Shape-based morphological filtering scheme: classical connected operators (black path) and connected operators on tree-based shape spaces (black+red path). Illustration from [Xu et al., 2016]

space $S_{\mathcal{T}}$ induced by a tree-based image representation \mathcal{T} . It is composed of three steps:

- tree \mathcal{TT} construction from tree-based image representation \mathcal{T} ;
- tree \mathcal{TT} filtering to obtain a simplified tree \mathcal{TT}' ;
- and a simplified tree \mathcal{T}' reconstruction from the simplified tree \mathcal{TT}' .

Let first tree \mathcal{T} be any tree-based image representation. Two tree based image representations are reviewed in Section 2.2.6. A tree-based shape space given by tree-based image representation \mathcal{T} with values assigned to nodes by an attribute function \mathcal{A} can be represented by a node-weighted graph $\mathcal{G}_{\mathcal{T}}$, where the nodes of graph are the nodes of the tree \mathcal{T} and the edges of graph $\mathcal{G}_{\mathcal{T}}$ are links created according to parenthood relationship between nodes. The second tree \mathcal{TT} is a component tree built on graph $\mathcal{G}_{\mathcal{T}}$, where the first tree \mathcal{T} was built from the input image. Based on the application and the characteristics of the attribute function \mathcal{A} , the second tree \mathcal{TT} can be chosen to be max tree or min tree. Considering the application object segmentation, when local minima of some attribute function \mathcal{A} corresponds to meaningful objects, the filtering on second tree \mathcal{TT} helps to remove meaningless local minima. In order to make the second tree filtering a simple pruning strategy, the second attribute \mathcal{AA} characterizing nodes of \mathcal{TT} is always designed to be increasing criterion. The step of simplified tree \mathcal{T}' reconstruction is trivial, which is achieved by removing the corresponding nodes $\{\mathcal{N}_i\}$ of tree \mathcal{T} contained in the series of filtered nodes \mathcal{NN}_k . This framework of shape-based morphology which is basically inspired from connected operators can be useful for shape filtering (connected filters on tree-based shape spaces) and object detection/segmentation. Several applications of this framework ranging from hierarchical image simplification, to shape filtering for retinal images, to object segmentation have been proposed in [Xu, 2013; Xu et al., 2016]. The application of shaping for PET image segmentation has been presented in the work of [Grossiord et al., 2015].

2.2.8 Challenges with connected filters: second order connectivity

As mentioned in Section 2.2.1 the important step in the general definition scheme of connected operators is transforming the image (or graph) into an equivalent hierarchical representation. And a major issue related to this step of hierarchical image representation is to make the structures of interest visible in this hierarchical representation. The role of connectivity is significant to solve this issue.

Until now, we implicitly considered the connected components and the flat zones on which connected operators work to be the usual four- or eight-connectivity for 2-D images. These connectivities relies on graph-based path connectivity: two pixels belong to same connected component or flat zone if they can be connected by a path. However, this approach to connectivity has many drawbacks. One of them being inability to remove boundary noise due to strict preservation of edges. Furthermore, objects broken up by, *e.g.*, sampling errors, are treated as different objects rather than part of the same. Conversely, disjoint real objects connected by noise pixels cannot be treated separately. More fundamentally, path connectivity may not correctly correspond to perception of human observer of an object as either distinct or single perceptual group. Many of the non-classical connectivities are obtained from path connectivity in some way. Second-generation connectivity [Ronse, 1998; Braga-Neto and Goutsias, 2002] is realized by means of a connectivity opening which is associated with a structural operator. The dependency on this operator may define constraints for the image domain to be connected. One of the most intuitive of these cases are extension of connectivity by means of a clustering operator, or so called clustering based connectivities [Braga-Neto and Goutsias, 2002]. These take the path-connected components of an image, and cluster them based on the size of separation between them. Thereby, the clusters of path-connected components now form the new connected components. In this case, any connected component in the base representation, is also connected in the new connectivity class, however the reverse is not valid.

To define a clustering-based connection, a simple approach is to apply an extensive operator Ψ such as dilation or a closing to the original image X with structuring elements. These operators will narrow the gap between connected components in the original image. The connected components are defined on this modified image, and then intersected with the original image. This means that the new connected components are defined on $\Psi(X)$. But as $\Psi(X)$ may be larger than X , it is required to compute the intersection of the connected components in $\Psi(X)$ and X . Alternatively, we can partition path-connected components in the original image into multiple fragments by cutting them, *e.g.*, narrow connecting bridges between wider regions. This idea leads to contraction-based connectivity described by [Ronse, 1998]. In this case, Ψ has to be an anti-extensive operator, like an opening by a connected structuring element.

In some cases, we might want to cluster in certain regions and partition in some other regions. This represents the case of mask-based connectivity introduced by [Ouzounis and Wilkinson, 2007a]. The idea here is to compute a mask image M

by any method and then use the connected components of this mask to modify the connectivity of the original image X . Therefore, instead of using an operator to modify the connectivity, a mask image is used as the modified image $\Psi(X)$ mentioned for clustering- or contraction- based connectivities. Instead, M and X can be images of the same scene obtained at different wavelengths (*e.g.* IR and visible in astronomy), or in different modalities (*e.g.* registered multi-modal medical image analysis).

These different types of second-order connectivities and their applications preserves the information of the original images, *i.e.*, these connectivities generate unique partition of the image domain. This property helps to reconstruct the image after filtering operations. However, in this work we present the concept of connected components in structural scale space built above the image scale space which relates closely to human perception of sparse objects, like empty catheter in fluoroscopic images described in Chapter 3.

Once the image has been transformed into equivalent hierarchical representation, one of the popular implementation of connected operators explained in Section 2.2.3 is image filtering and segmentation. This also corresponds to the second issue with connected operators of distinguishing structures of interest in the hierarchical representation. The task of image filtering involves designing of attribute function \mathcal{A} as described in Section 2.2.3.3, a filtering of hierarchy which simplifies it based on attribute function \mathcal{A} and image reconstruction from the simplified hierarchy that consist the filtering result [Salembier and Garrido, 2000; Salembier and Wilkinson, 2009]. The attribute function \mathcal{A} can also be designed to segment a particular region of the image. We previously discuss these image filtering and segmentation techniques in Section 2.2.7.

2.3 Registration

Image registration, often referred as image fusion, warping or matching, can be defined as process of finding the transformation between two or more images so that their structures of interest align. Medical image registration is so vast topic and so widely studied for various applications and modalities. Registration methods and algorithms can be classified in different ways and in literature several reviews have been published compose them together [Oliveira and Tavares, 2014; Zitova and Flusser, 2003; Maintz and Viergever, 1998]. We restrict the goal of this section to understand the context of image registration with respect to the vessel of intervention detection problem to be solved in this work.

Image registration is a crucial step for image analysis in which valuable information is conveyed in two images; *i.e.* images acquired at different times and/or from different view points and/or in different modalities (different sensors). This information conveyance is essentially represented by a transformation that compensates the absence of geometrical information during the acquisition of two images to be registered. The registration problem can be well defined as finding the “optimal” transformation that “best” aligns the two considered images. Here, the adjective “optimal” for transformation donates a sense of a minimization problem and the

adjective “best” for alignment hints a distance or dissimilarity measure that is minimized. Thus, a general formulation of registration problem is as a solution to equation:

$$\hat{T} = \underset{T \in \Omega}{\operatorname{argmin}} \mathcal{D}(T(g), f) \quad (2.17)$$

where \hat{T} is the desired optimal transformation between two images g and f , \mathcal{D} is called as objective function (or similarity measure) quantifying the distance or dissimilarity that is minimized and Ω is set of admissible transformations. Based on the information represented, the similarity measures are divided into two classes:

- Intensity-based similarity measures
- Feature-based similarity measures

Depending on features used some similarity measures can be included in both classes. Intensity-based measures use all image pixels and do not require any feature extraction. They tend to be more stable near the minima of the objective function as they do not depend on uncertain feature locations or on correspondences which may fluctuate with slight changes in transformation. They are used in cases when the structures to be matched are not sparse in image or challenging to segment. They are based on intensity differences, intensity cross-correlation or information theory (mutual information). On the other hand, feature-based measures tend to be faster, have wide capture range and allow registration quality to be focused only on selected (sparse) subset of information in both images. The selected sparse subset of image information essentially refers to the features which are segmentations of the structures of interest in the images. Thus, these methods strongly depend on the segmentation accuracy. The segmentations representing the features can be set of non-connected points, open/closed curves or more complex structures like trees or graphs. The strengths and weaknesses of the two measures are studied with a quantitative approach in [McLaughlin et al., 2001].

The task of vessel of intervention detection is about determining the position of guidewire tip in the vasculature from two fluoroscopic images: one obtained during contrast agent injection in the vasculature and another during navigation of guidewire when vasculature is not visible. An easy approach would be extracting the features from the two images, *i.e.* the vessels modeled by their centerlines and segmented guidewire tip, and find the position of the guidewire tip in the segmented vessel. This task can be modeled as a feature-based registration or matching problem, where we solve the Equation 2.17 by minimizing the distance \mathcal{D} between two geometric representations (features). The vessel centerlines and the segmented guidewire tip can be represented by different types of features. The simplest feature representation can be a set of points representing the centerlines of the vessels and the centerline of the segmented guidewire tip [Kim et al., 2017]. In this work, the registration is conducted using the Iterative Closest Point (ICP) algorithm, first proposed by [Besl and McKay, 1992]. The ICP algorithm consist of two steps: 1) pair each point of model to its closest in the data; 2) eliminate the transformation minimizing the distance between paired points. The closest point pairing assumption in the ICP framework can lead to incorrect pairings causing some part

of guidewire tip to pair to one vessel and some other part of guidewire tip to pair to other vessel. Similar issue in case of 3D/2D registration has been mentioned in the work of [Kita et al., 1998]. Though we can relate the task to guidewire tip to vasculature registration to that of 3D/2D vascular registration and take aid of the corresponding vast literature, however there are significant major differences between the two problems: 1) guidewire tip is small as compared to vasculature; 2) thereby, a non-rigid registration can lead to multiple pairings of guidewire tip to the vasculature at different positions.

Taking into account these anomalies as compared to 3D/2D registration, another natural way of representing the guidewire tip and vasculature can be a curve and set of curves respectively. A curve can be a set of ordered points, a continuous interpolation passing through them. In the work of [Liu et al., 1998], 3D and 2D structures were represented as curves and a new point pairing strategy was proposed using the local vessel direction to restrict the search zone for point pairing candidates. This is followed by optimization of sum of squared Euclidean distance using the Newton approach. The work of [Benseghir et al., 2015] was an advancement in registration algorithm endorsing the particular structure of the vasculature to be matched. The proposed Iterative Closest Curve (ICC) framework for registration of curvilinear structures allows curve pairing by avoiding to penalize incoherence in the point pairing set by imposing coherence using a dedicated distance called Fréchet distance. In case of vessel of intervention detection problem, the curve belonging to guidewire tip can be paired to its corresponding position in the vasculature using the curve pairing imposed with Fréchet distance. More details on this curve pairing are discussed in the Chapter 4.

Empty catheter segmentation

Contents

3.1	Abstract	51
3.2	Introduction	52
3.2.1	Guiding catheter vs other catheters	52
3.2.2	Related works	53
3.2.3	Guiding catheter in fluoroscopy	54
3.2.4	Proposed method at a glance	56
3.3	Scale spaces and deep connected components	56
3.4	Method	58
3.4.1	Curve blobs extraction	58
3.4.2	Curve blob clustering	60
3.4.3	Deep connected component selection	64
3.5	Segmentation Quality Evaluation	65
3.5.1	Dataset	65
3.5.2	Ground truth construction	66
3.5.3	Segmentation	66
3.5.4	Evaluation measures	66
3.6	Results and discussion	69
3.7	Conclusion of catheter segmentation	73

3.1 Abstract

In this chapter, we present a method for segmentation of empty guiding catheter in fluoroscopic X-ray images. The guiding catheter, being a commonly visible landmark, its segmentation is an important and a difficult brick for PCI procedure modeling. In number of clinical situations, the catheter is empty and appears as

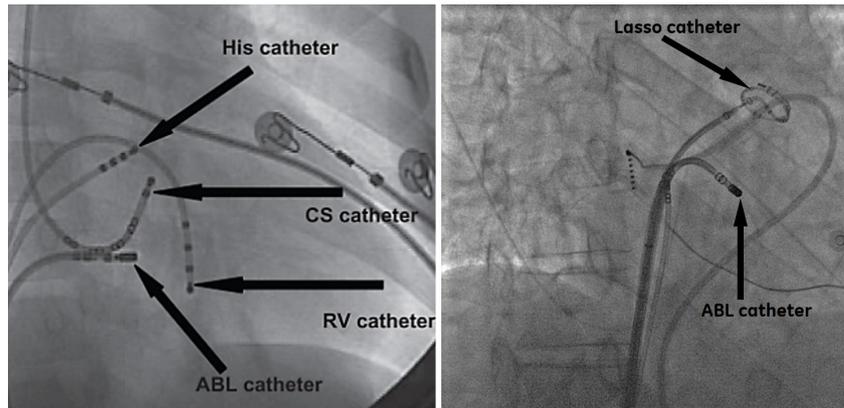
a low contrasted structure with two parallel and partially disconnected edges. To segment it, we work on the level-set scale-space of image, the min tree, to extract curve blobs. We then propose a novel structural scale-space, a hierarchy built on these curve blobs. The deep connected component, *i.e.* the cluster of curve blobs on this hierarchy, that maximizes the likelihood to be an empty catheter is retained as final segmentation. We evaluate the performance of the algorithm on a database of 1279 fluoroscopic images from 6 patients. As a result, we obtain very good qualitative and quantitative segmentation performance, with mean precision and recall of 83.85% and 67.87% respectively. We develop a novel structural scale-space to segment a structured object, the empty catheter, in challenging situations where the information content is very sparse in the images. Fully-automatic empty catheter segmentation in X-ray fluoroscopic images is an important and preliminary step in PCI procedure modeling, as it aids in tagging the arrival and removal location of other interventional tools.

3.2 Introduction

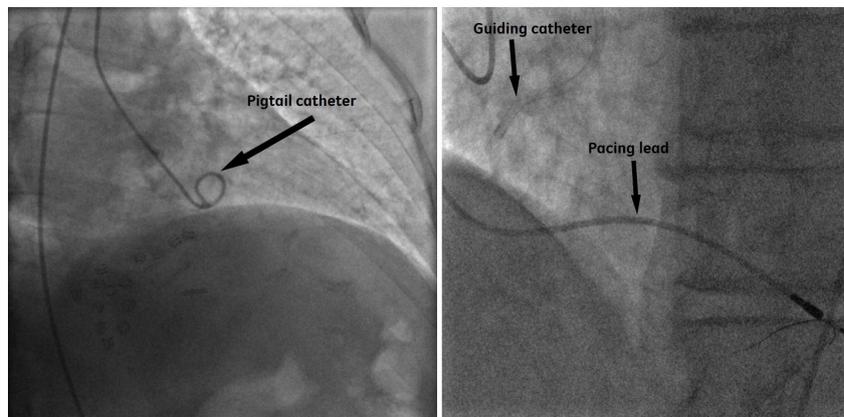
3.2.1 Guiding catheter vs other catheters

A catheter is a flexible surgical tube or instrument inserted in body cavity or vessel. In cardiology, numerous types of catheters are used in X-ray guided electrophysiology (EP) and PCI procedures. In X-ray images catheters appear as tubular structures with low contrast. These catheters can be mainly distinguished by their shape and features of their tip. Various EP catheters and the catheters used in PCI procedures are shown in Figure 3.1. EP catheters are used for electrical mapping and ablation and can be characterized by features such as electrodes which are well contrasted in X-ray images. The different types of EP catheters are radio-frequency (RF) ablation catheter, circumferential mapping (or lasso) catheter, coronary sinus catheter, right ventricle (RV) catheter, His catheter. The catheters used in PCI procedures are guiding (Judkins) catheter and pigtail catheter. Guiding catheter, which is our tool of interest here, is used throughout PCI procedures for performing various tasks. Whereas, pigtail catheters are used for calculating ejection fraction. Figure 3.1b also contains pacing lead, which may appear in X-ray images during PCI procedures and looks like a catheter.

A guiding catheter is a tool that appears throughout the PCI procedure. It can contribute to significant semantic information since it is the first tool to appear in the field of view. The guiding catheter is fixed at ostia of left or right coronary artery for rest of the procedure. It is the conduct for all other tools/devices which will be brought into the vessel where the lesion is present. The contrast agent is also delivered into the vasculature by injecting it through the guiding catheter. Thus, the segmentation of guiding catheter can help in procedure modeling to determine the events/phases of the arrival and removal location of other devices (guide wire, marker balls).



(a) EP catheters



(b) PCI catheters

Figure 3.1: Different types of catheters in cardiac fluoroscopy: (a) Electrophysiology (EP) catheters, abbreviations: ABL catheter, ablation catheter; CS catheter, coronary sinus catheter; RV catheter, right ventricle catheter; His catheter, (b) two types of PCI catheters including pigtail catheter and guiding catheter. Pacing lead, a tubular tool, also looks like a catheter.

3.2.2 Related works

In literature, most of the approaches focusing on catheter segmentation and tracking in X-ray images are catheter type-specific or catheter tip type-specific and cannot be extended to other types. These works have particularly addressed the segmentation and tracking of electrophysiology (EP) catheters. EP electrodes are clearly visible and well contrasted, which can be localized to segment the EP catheters. For segmentation and tracking of coronary sinus catheter tips, [Ma et al., 2010] proposed electrode-based detector, which identifies blob-like structures based on Hessian matrix analysis. [Wu et al., 2011] proposed learning based approach for CS catheter segmentation consisting of a blob-detector trained with probabilistic boosting trees (PBT). [Milletari et al., 2013] showed that it's possible to automatically detect multiple and overlapping CS catheters with segmentation accuracy of 99.3%. This method was based on blob detection and clustering of CS electrodes.

The electrode-based detector of [Ma et al., 2010] was extended for detection of ablation catheter and lasso catheters [Ma et al., 2013]. [Brost et al., 2009, 2010] has also proposed efficient tracking of lasso catheter for electrophysiology procedures. Though the performance of segmentation of EP catheters evolved with state-of-the-art methods [Milletari et al., 2014; Wu et al., 2015; Baur et al., 2016], these methods are all based on the well contrasted electrodes of EP catheters. Similarly, radio-opaque markers of the pigtail catheters were used to design a dedicated method for its segmentation by [Lessard et al., 2015].

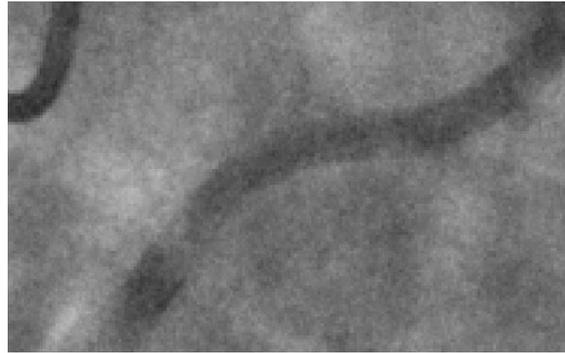
Most of these approaches for catheter segmentation and tracking are specific to the catheter types and catheter tip types. Hence, these approaches cannot be extended segmentation of other types of catheters. Segmentation of catheters in absence of salient features like electrodes or tips has been studied less frequently. There are very few approaches that focuses on tracking of whole catheter. Some of these methods [Baert et al., 2003] enhance the catheter and guidewire with the help of Hessian-based filters followed by spline fitting approach. Many approaches model the catheter as B-spline, which is parametric representation of non-linear curve. Other popular approaches include modeling tools with B-splines whose control points are determines with discrete optimization [Heibel et al., 2013] and B-spline tube models [Chang et al., 2016]. [Chen et al., 2016] proposed guidewire segmentation with B-spline refinement on sequence of small segments followed by guidewire tracking which was formulated as graph-based optimization problem. [Wang et al., 2009] proposed segmentation of guidewire between the catheter tip and guidewire tip.

The major drawbacks of these methods are: i) the first frame of the fluoroscopic sequence has to be annotated manually and ii) the length and the curvature of the guidewire and catheter should not change much during the sequence. To overcome these challenges, [Wagner et al., 2016] proposed a fully automatic method using directional noise reduction and path extraction. Recently, [Ambrosini et al., 2017] proposed deep convolutional neural network based faster and fully automatic segmentation of guidewire and catheter. All these approaches focus on tracking of guidewire along with the catheter body.

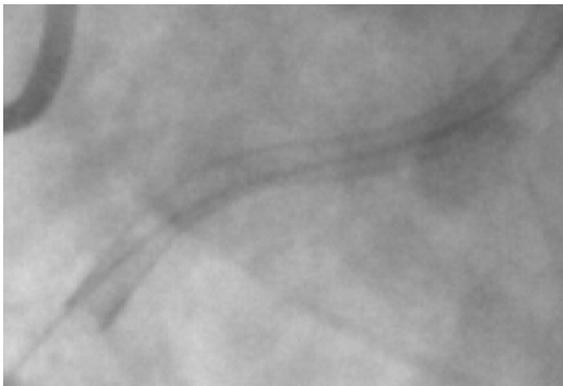
As underlined previously in Section 1.3, the workflow of the procedures shall not be modified by requiring additional user interactions. So the approach shall not only include the capability to track the catheter from image to image but also the capability to locate it without any type of initialization. This is why we propose to focus our work on the task of searching the empty catheter in individual X-ray image. In a second step, temporal tracking strategies can be used to increase the robustness, eliminate outliers and focus on catheter tip. In this work, we address the task of empty catheter segmentation in fluoroscopic images.

3.2.3 Guiding catheter in fluoroscopy

A guiding catheter is a tool that appears throughout the PCI procedure. We address here the empty catheter case i.e. when it is not filled with contrast media. Such empty catheter appears in 30% to 40% of the images acquired during PCI procedure



(a) Filled catheter



(b) Empty catheter in record image



(c) Empty catheter in fluoroscopic image

Figure 3.2: Guiding catheter in fluoroscopy: (a) filled catheter, empty catheter appearance in (b) record image and (c) fluoroscopic image. All other components of the imaging situation (patient, geometry) are identical. Fluoroscopic images are noisier and the contrasts are weaker than in the record images

and mostly in the first steps of the procedure where the analysis shall start. A filled catheter (with contrast agent) is highly contrasted structure and relatively easy to segment. Figure 3.2 shows the difference between the filled and empty catheter. From first analysis, we observe that the segmentation of empty guiding catheter is of utmost importance.

The appearance of empty guiding catheter also varies according to the imaging modes of the interventional systems. Figures 3.2b and 3.2c, shows a record and a fluoroscopic image respectively, taken at few seconds of interval in the same setting illustrate the difference of quality between these two imaging modes of interventional angiographic units. However, the segmentation task is difficult due to low contrast in fluoroscopic images. In these images the empty catheter appears as a low contrasted structure with two parallel and partially disconnected edges because it is just an empty tubular pipe made of a material with little radio-opacity. As X-ray contrast of the object depends on both the radio-opacity of the material and its thickness, an empty catheter is mainly detectable on its boundaries, where the projective thickness is larger. So overall the image signal can be characterized by a general geometric structure coming from smooth curve of the catheter and sparse

information due to the limitations of X-ray imaging with low dose.

3.2.4 Proposed method at a glance

We devise a bottom up approach for segmenting the empty catheter in fluoroscopic images. We first use the level-set scale-space, *i.e.*, the hierarchy of all the level sets of the gray scale image, called the component tree, to extract *curve blobs*, small dark persistent regions that are potentially part of the empty catheter. These curve blobs are disconnected in the image space. We then propose a structural graph-based scale-space, in the form of a hierarchy (*i.e.*, a tree), where these curve blobs are connected. We analyze this hierarchy to select the cluster of curve blobs that maximizes a score of likelihood to be an empty catheter. If the first tree exhibits the deep structures of the critical points, the second tree puts forward the even deeper structures of interest, that we call *deep connected components*. To evaluate our work, we use a database of 1279 fluoroscopic images from 6 patients (retrospective use of collected images of patients). The centerline of the catheter in this dataset was manually delineated by a trained observer to define the ground truth.

3.3 Scale spaces and deep connected components

Classical techniques [Frangi et al., 1998] which are mostly differential-based, do not work in this situation due to weak contrast of empty catheters and high noise level of fluoroscopic images. We decided to adopt an approach derived from the theory of scale-space. According to this theory, each structure in a scene is visible at a certain scale. Finding the right scale is challenging issue that has been studied by many authors, primarily using the Gaussian scale-space. [Lindeberg, 1993] studied the problem of linking local critical points (extrema and saddle) over scales, leading to the so-called scale-space primal sketch which makes explicit the relation between structures at different scales. An important practical issue in this approach is the ability to attach a persistence measure to the structures, *i.e.*, a measure of the duration of survival of the structures during the evolution. In their seminal works, both [Koenderink, 1984] and [Witkin, 1984] propose to investigate the *deep structure* of an image, *i.e.*, the structure of all levels of resolution simultaneously.

For signal of dimension 2 or greater, two drawbacks of the Gaussian scale-space are that, during the evolution, 1) structures evolve (change shape), and 2) critical points can be created. On the other hand, connected filters from mathematical morphology [Salembier and Wilkinson, 2009] can be seen as a non-linear scale space: in such approaches, the image is transformed into an equivalent tree-based representation (tree of upper-level sets, of lower level sets, or both), and attributes can be computed for each node of the tree. Selecting the nodes with a criterion based on these attributes allows to study the evolution of the nodes of the tree, and in particular their persistence. Obviously, during such evolution, structures cannot change, and no novel structure can be created. A formalization of such ideas in the context of image segmentation has been achieved by [Guigues et al., 2006]. The hierarchical data organization presented in this chapter has the main scale-space

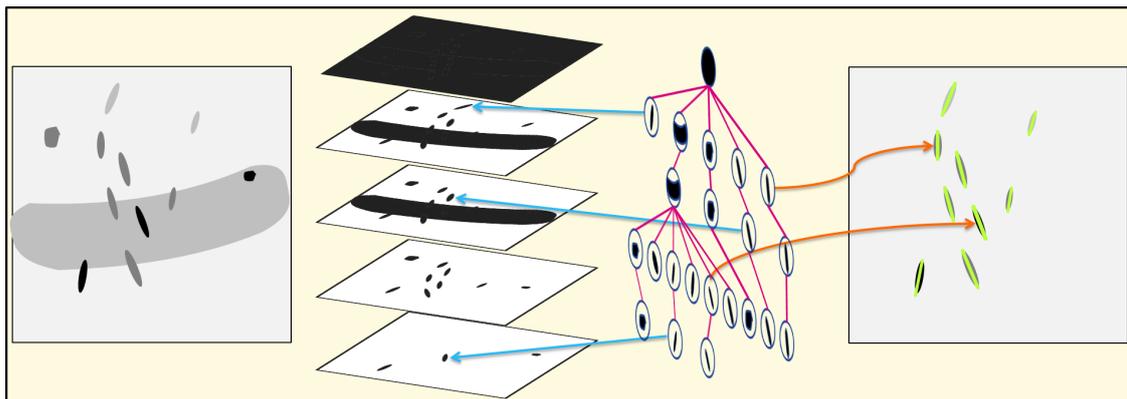


Figure 3.3: Curve blobs extraction with min tree filtering, from left to right: Input toy image; lower level-sets of input image; corresponding min tree where each node is one of the connected components of the level-sets; extracted curve blobs overlaid with green centerlines.

properties studied by [Guigues et al., 2006]. Early attempt for using hierarchical data organization for guidewire localization has been made by [Barbu et al., 2007]. They use marginal space learning based hierarchical model of curves (obtained from low-level segment detector) to model complex free-form curves. The similarity with our approach is that both are bottom up approaches with low level segment/blob detector as first step. Though [Barbu et al., 2007] did not show segmentation of empty catheter, a head-to-head comparison would be helpful but neither the dataset nor the implementation has been made public. Reported computational time are close to ours. A major insight that we draw from such methods is that any hierarchical data organization has the main scale-space properties. As the algorithms for computing the trees are graph-based ones, these ideas can be extended to work on any graph, and not only on 2D/3D images (*e.g.* [Xu et al., 2016], [Najman and Cousty, 2014]).

In this work, we first use the level-set scale space to identify curve blobs, which are small dark persistent regions that are potentially part of the empty catheter. We propose a novel structural graph-based scale-space, in the form of a hierarchy, *i.e.* a tree built on the curve blobs. We analyze this second tree with the very same techniques as the first one, and we retain the most persistent structures in this second scale-space as the final segmentation. If the first tree exhibits the deep structures of the critical points, the second tree puts forward the even deeper structures of interest, that we call deep connected components.

The major steps of our proposed algorithm are detailed in section 3.4. The strategy retained for assessing the performance of this algorithm is discussed in the section 3.5. The qualitative and quantitative results are presented in section 3.6. The main contributions are:

- proposal of a fully automatic catheter segmentation method in absence of guidewire and contrast agent;
- proposal of a new notion of a deep connected component, appearing in a

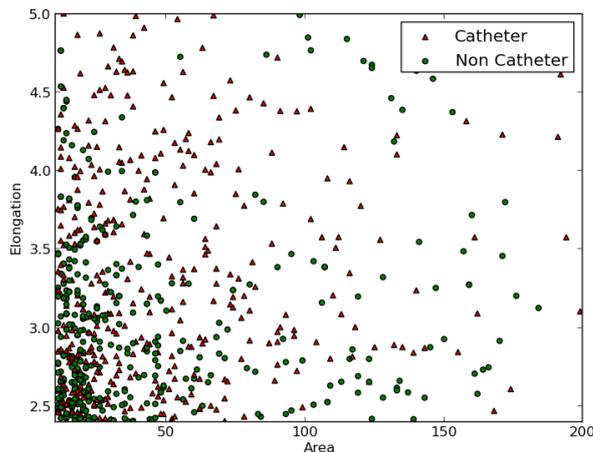


Figure 3.4: Distribution of connected components belonging to two classes (catheter and non-catheter) in the attribute space of area and elongation.

second-order scale-space (Section 3.4); and

- assessment of the proposed method on a database of 1250 fluoroscopic images (Section 3.6).

3.4 Method

Our bottom-up approach comprises two main parts: the first one aims at identifying small dark regions called curve blobs (Section 3.4.1) and the second one focuses on grouping them in order to retrieve the whole catheter (Section 3.4.2 and Section 3.4.3). Both steps are performed by analyzing a hierarchical structure, called a component tree [Salembier and Wilkinson, 2009].

3.4.1 Curve blobs extraction

Curve blobs are small dark persistent regions that are potentially part of the parallel edges of the catheter. Various methods in literature based on Haar or ridge features can be used for extraction of curve blobs. We are interested to extract the curve blobs from the image using a component tree called min tree. As described previously in Section 2.2.7 a unique framework of min tree can be used for segmentation of different interventional tools by designing dedicated filtering strategies for each interventional tool.

As a first step, we pre-process the input image to remove the anatomical background with the help of morphological operation. The image f on which the min tree is computed is obtained with a simple morphological dark top-hat on the input X-ray image with a circular structuring element whose radius is same as radius of catheter. According to the min tree framework (Section 2.2.6), for a grayscale image

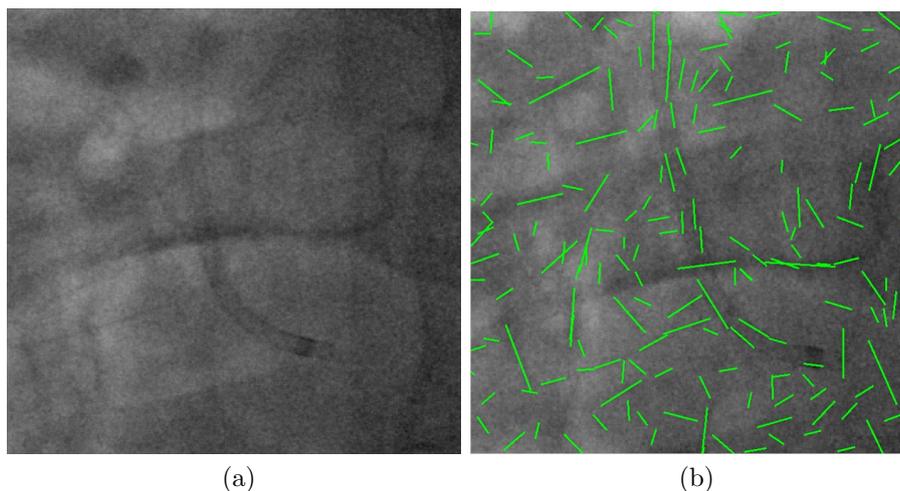


Figure 3.5: Curve blob extraction, (a) input fluoroscopic image and (b) green centerlines of extracted curve blobs.

(pre-processed fluoroscopic image in this case) f defined on a domain \mathcal{D} and with values in range \mathbb{R} the *lower-level set of f with level h* in \mathbb{R} is

$$f_h = \{x \in \mathcal{D} \mid f(x) \leq h\} . \quad (3.1)$$

The image f when thresholded in an increasing order at every possible gray level in the range \mathbb{R} , yields a stack of nested (lower-) level sets. Each level set can be partitioned into connected components when the domain \mathcal{D} is structured as a pixel-adjacency graph (we consider 4-adjacency relation). Based on inclusion relationship, the structured set of all connected components is a directed tree called the min tree of the image f , denoted by \mathcal{T}_f . These connected components are the nodes of the tree \mathcal{T}_f , and are denoted by \mathcal{N} .

The min tree considers only the dark connected regions of the image and the curve blobs appear as regions at different levels of this tree (*i.e.*, at different scales, refer Figure 3.3 for illustration). Thus, two curve blobs might well be obtained from two distinct threshold values. Then, the set of all curve blobs is included in a non-local (*i.e.*, a spatially variable) threshold of the image, that is a non-horizontal cut of the min tree. In order to obtain curve blobs among the connected components, we design a criterion that selects them in this non-horizontal cut. We assign to any component in the min tree, attributes characterizing its shape and structural properties. For curve blobs, we design a selection criterion based on two attributes:

- **Area:** The area attribute refers to the number of pixels in the connected component.
- **Elongation:** The elongation is given by $1 - l_{min}/l_{max}$, where l_{max} and l_{min} are the lengths of the axes of an ellipse optimally fitted to the component.

In order to select curve blobs, a straightforward idea is to select components whose area is in a certain range and with elongation attribute large enough. Note

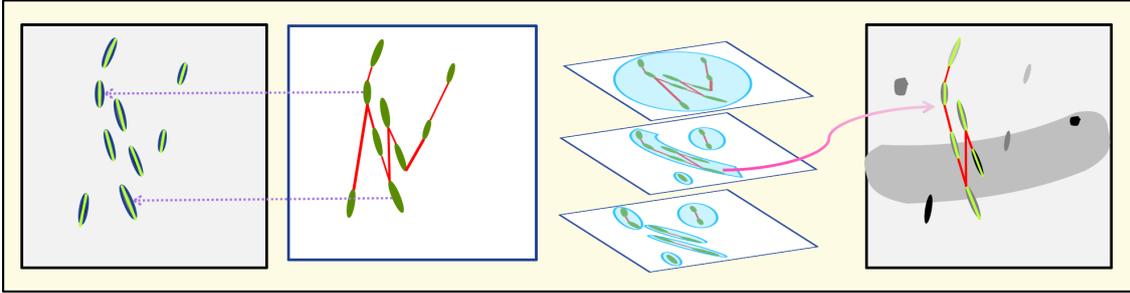


Figure 3.6: Curve blobs clustering; from left to right: extracted curve blobs (see Figure 3.3); connection of curve blobs on structural scale-space; hierarchy of deep connected components (clusters of curve blobs at different scales); selected deep connected component overlaid on input image.

that we are not interested in too small connected components which might come from noise nor in too large ones which might correspond to filled catheters, pacing leads, or other large anatomical or interventional structures. To establish a proper selection criterion, we built a set of curve blobs belonging to empty catheter (taken in 14 images) and another set of curve blobs selected randomly in the same image at location away from the marked empty catheter. We investigated the distribution of these two sets, catheter and non-catheter blobs, in the area-elongation space (see Figure 3.4). We observed that the two classes of catheter and non-catheter blobs appear together in this space. We established a relevant criterion: defined by independent lower limit for area and elongation and a maximum upper limit on the weighted sum of area and elongation. All nodes satisfying this criterion are selected to form the set \mathcal{C} of curve blobs. Sometimes nested connected components of min tree satisfy the selection criterion as they depict same region in the image. A filtering on the component tree, \mathcal{T}_f , is performed in order to preserve elements with largest area (taking aid of the inclusion relationship). Thus we propose a strategy to preserve the parent node \mathcal{N}_p of node \mathcal{N} ,

$$\text{If } \mathcal{N}, \mathcal{N}_p \in \mathcal{C} \mid \forall \mathcal{N} \in \mathcal{T}_f \quad (3.2)$$

prune the children nodes \mathcal{N} . Figure 3.3 illustrates with an toy example, the extracted curve blobs with these filtering steps. As evident from the distribution in the area-elongation space, the number of curve blobs in fluoroscopic image can be significantly large. The extracted curve blobs with fluoroscopic image as input are shown in Figure 3.5.

3.4.2 Curve blob clustering

This section presents the main idea of empty catheter detection, *i.e.* curve blob clustering in the structural scale space. Fig. 3.6 intuitively portrays this idea of curve blob clustering in the structural scale space. Some curve blobs extracted in the previous step are regions of edges of the catheter, while some others correspond to other anatomical and interventional structures or to noise. By analyzing indi-

vidually a given curve blob, it is difficult to decide whether it is part of a catheter because contextual information is missing. So we consider them in a common space and define a weight for each pair of curve blobs, called as *blob pair weight*. The weight is defined to be small when the two considered curve blobs are likely to be part of an empty catheter. In rightmost image in Fig. 3.6, the selected cluster of curve blob belonging to empty catheter is shown where green curve blobs are connected with red edges which link blob pairs. We propose to build it by combining three elementary weights, each of them characterizing one aspect of the relation between two curve blobs. Let us consider C_1 and C_2 be any two curve blobs and their blob axes (Ca_1 and Ca_2 respectively) are the axes of the optimally fitted ellipses to the respective blobs. These blob axes represent the orientations (or first moments) of the blobs. Let b_1 and b_2 be the barycenters of two blobs and \vec{b} be the vector joining these barycenters as in Figure 3.7. The three elementary weights are defined as,

- **Spatial weight**, w_S : The intent of the hierarchy is to connect primarily the curve blobs which are sufficiently close. So we include a spatial weight, which is the Euclidean distance between two blobs. To be precise, the Euclidean distance between two blobs is the minimum distance between the blob axes of the pair of blobs considered.
- **Alignment weight**, w_A : In the image space, empty catheter looks like partially disconnected elements of curvilinear structure. Thus, the blobs belonging to an empty catheter shall be a part of a smooth curvilinear structure and as such shall have a close orientation when they are not distant from each other. The alignment weight is designed to measure this property. Two considered blobs with their barycenters are enriched by orientation vectors $\vec{Ca_1}$ and $\vec{Ca_2}$ representing the blob axes. Thus, the alignment weight of two blobs is derived using the minimum of the inner product of these orientation vectors and of vectors joining barycenters as given by,

$$w_A = 1 - \left| \min(\langle \vec{Ca_1}, \vec{b} \rangle, \langle \vec{Ca_2}, \vec{b} \rangle) \right| \quad (3.3)$$

- **Profile weight**, w_P : We have mentioned that, in an image, empty catheter's appearance reduces to two parallel edges. In order to incorporate this information in the strength measure, we need to perform a measurement on the intensity of the blob and its immediate neighborhood in the image. Profile weight estimates the dissimilarity between the intensity profiles along a pair of blobs and the desired intensity profile along an ideal empty catheter. The intensity profile \mathcal{P} along a blob is a series of values ($p[d] \mid d \in [-N, N]$), where $p[d]$ is the average of the intensities of the pixels on the segments parallel to the blob axis located at (signed) distance d . The intensity profile along a blob on the edge of catheter should ideally have two minima corresponding to the two parallel edges of the catheter. Let the intensity profiles of given two blobs C_1 and C_2 be \mathcal{P}_1 and \mathcal{P}_2 as shown in Figure 3.7. We compute

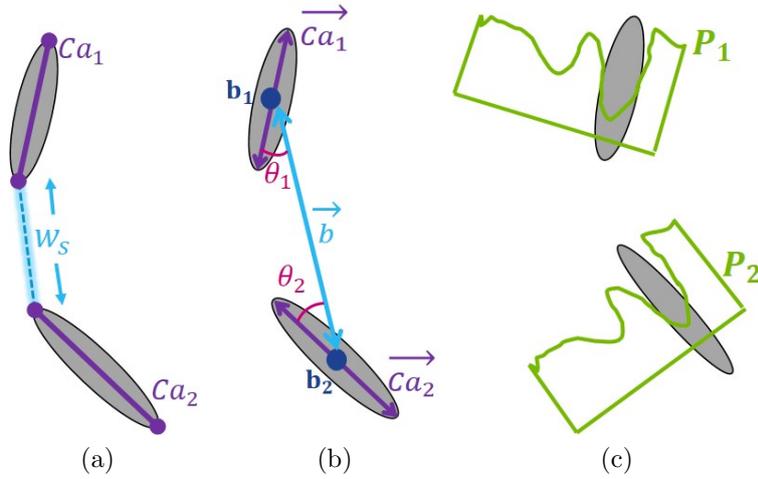


Figure 3.7: Blob pair weight computation, (a) w_S is the spatial weight between two blobs with blob axes Ca_1 and Ca_2 (b) vectors to compute alignment weight and (c) shows profiles \mathcal{P}_1 and \mathcal{P}_2 to compute profile weight.

the expected intensity profile \mathcal{P}_{cat} as the mean intensity profile of the blobs belonging to catheter in the ground truth images. The distance to expected intensity profile of a blob is computed as,

$$\delta_P(C_n) = \sum_{d=0}^{2N} |\mathcal{P}_n[d] - \mathcal{P}_{cat}[d]| . \quad (3.4)$$

The profile weight of a blob pair is the mean of the distances to expected profile of the two blobs,

$$w_P = \frac{\delta_P(C_1) + \delta_P(C_2)}{2} \quad (3.5)$$

We transform the individual weights (w) with sigmoidal function,

$$w^{\sim} = \frac{1}{1 + e^{-k(w-w_o)}} \quad (3.6)$$

so that the three weights lie in the same range. Thus we obtain w_S^{\sim} , w_A^{\sim} and w_P^{\sim} by establishing the parameters k and w_o of Equation 3.6 for each of the three weights from simple numerical consideration. The *blob pair weight* w_B is computed for any two curve blobs C_1 and C_2 in \mathcal{C} as:

$$w_B(C_1, C_2) = \alpha * w_S^{\sim}(C_1, C_2) + \beta * w_A^{\sim}(C_1, C_2) + (1 - \alpha - \beta) * w_P^{\sim}(C_1, C_2) \quad (3.7)$$

The hierarchy of curve blobs can be defined with using the blob pair weight. Intuitively, a threshold on the blob pair weight gives a partition of \mathcal{C} into clusters of connected blobs which are “consistent” with respect to blob pair weights. More

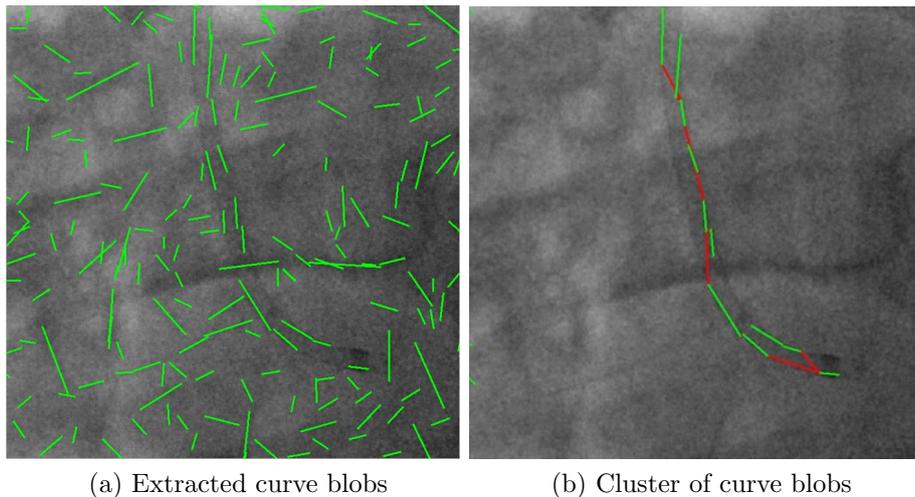


Figure 3.8: Curve blob clustering, (a) green centerlines of extracted curve blobs with the input image (Figure 3.5a); (b) selected cluster of curve blob representing empty catheter.

precisely, for a given threshold value λ , we build a curve blobs graph $G_\lambda = (\mathcal{C}, E_\lambda)$ where each vertex is a curve blob in \mathcal{C} and where two curve blobs are linked by an edge in E_λ if their blob pair weight is below λ :

$$E_\lambda = \{\{C_1, C_2\} \mid w_B(C_1, C_2) \leq \lambda, C_1, C_2 \in \mathcal{C}\}. \quad (3.8)$$

Such graph induces a partition \mathcal{P}_λ of the curve blobs into connected components, each element being referred to as *blob clusters at scale λ* . The set of all blob clusters obtained at every possible scale is a hierarchy \mathcal{H} of partitions on the set of curve blobs, given by:

$$\mathcal{H} = \bigcup \{\mathcal{P}_\lambda \mid \lambda \in [0, 1]\}. \quad (3.9)$$

Indeed \mathcal{H} is a hierarchy since any two blob clusters in \mathcal{H} are either nested or disjoint. Hence, this hierarchy can be managed as a tree structure where the parenthood relationship is given by inclusion relationship on the set of clusters (more precisely it is its Hasse diagram). Using the terminology from mathematical morphology, where a similar construction has been done for pixels with a different measure [Soille, 2008; Cousty et al., 2015], we name this precise component tree the quasi-flat zone hierarchy of the blob pair weight. Any element of a partition at scale λ is called a (quasi-flat) zone. Hence, in the context of this chapter, these zones refer to clusters of the curve blobs or *deep connected components*. This hierarchy is what we mention previously as a structural scale-space in which we are looking for the deep connected component corresponding to the empty catheter in the image. Unlike the min tree, which was directly built over the image pixels, at every threshold value, the set of all curve blobs is partitioned, the elements of the partitions being clusters of curve blobs.

3.4.3 Deep connected component selection

An empty catheter appears to be a zone (deep connected component) in the partition of this quasi-flat zones hierarchy at some scale λ in the structural scale-space. The object (zone) of interest here being the catheter which is elongated curvilinear object, its length can be best determined by the centerline. Determining the length of the zone is not straightforward. Zones are the connected components whose vertices are the curve blobs and the edges are the segments joining any two curve blobs. As a result, towards the analysis of the size and shape of zones (clusters of curve blobs) in the structural scale-space, we decide to fit a curve to each of the zones to determine their centerlines. As the zone belonging to the catheter may contain curve blobs from either one or both edges of the catheter (see Figure 3.8b), spline interpolation may not be best estimate for the centerline of the catheter. We go for a global fitting with a 3rd order polynomial curve for each zone. The order of the polynomial is determined on the basis of the possible appearance of the catheter in the fluoroscopic images. The data points considered to effectively perform polynomial least square fitting are the barycenter of the blob and the two ends of the blob axis for all the blobs in the zone. This classical fitting is done by minimizing a least-square error obtained by L_2 -norm of the difference between the blob (observed) points and fitted points.

In order to further analyze the shape of the zones, we have a measure \mathcal{L} which maps to any zone \mathcal{Z} a positive real value $\mathcal{L}(\mathcal{Z})$ that represents the likelihood of \mathcal{Z} of being an empty catheter. For each zone \mathcal{Z} the measure $\mathcal{L}(\mathcal{Z})$ depends on some attributes of zone \mathcal{Z} . These attributes depend on several geometric and (time and space) continuity properties of \mathcal{Z} modeling the appearance of empty catheter in fluoroscopic images sequences. The six zone attributes used in this study are explained below:

- *Length*: the length of a zone is the length of the 3rd order polynomial curve fitted on the zone as described above. Once the fitting is done, the segment of the curve is selected by mapping the blob points on the fitted curve and determining the extremities of these mappings on the curve. The length is the arc length of the curve between these extremities.
- *Fitting error*: the fitting error is calculated as average of the residual errors in the least square fitting above. For the zone belonging to catheter expected fitting error is small as curvilinear structure belonging to catheter can be fitted with 3rd order polynomial. However, when the zones are more spread in the space (not related to elongated object), the fitting error is larger.
- *Average distance to expected profile*: zone belonging to empty catheter should have information related to the parallel edges of the catheter. In previous steps, this information was imbibed in the profile weight where the distance to expected profile $\delta_P(C_n)$ for any blob C_n is defined (Equation 3.4). This attribute is the mean of distances to expected profile of all curve blobs of the given zone.

- *Proximity to expected scale of observation*: an empty catheter is expected to appear as a zone at a certain scale λ^* . Therefore, we design a proximity attribute to this expected scale. The expected scale is observed from the image dataset and is a value in the range $[0, 1]$, (refer Equation 3.9).
- *Proximity to image borders*: one end of the catheter in field of view always appears at the borders of the image frame. We incorporate this information to design an attribute by computing the minimum of the distances of all the curve blobs in a zone to the border of image.
- *Temporal feedback*: for each frame, we compute a feedback image with the score of the detected catheter in its bounding box and 0 elsewhere. We then sum the feedback images of the past 10 frames and consider the average for all pixels of a zone. It is taken as the temporal feedback attribute of the zone. We choose bounding box based method for temporal feedback because of the motion of the catheter across the sequence induced by heartbeat and breathing.

However, to homogenize the range of the attributes before combining them, we form homogenized scores as the images of the attributes by Gaussian functions whose parameters (centers and width) are determined by analyzing the ground truth (see Section 3.5 for details on ground truth). The measure \mathcal{L} is product of the six homogenized scores. The resulting segmentation \mathcal{S} is the zone in the hierarchy \mathcal{H} that maximizes the likelihood score:

$$\mathcal{S} = \arg \max_{\mathcal{Z} \in \mathcal{H}} \mathcal{L}(\mathcal{Z}) \quad (3.10)$$

3.5 Segmentation Quality Evaluation

For evaluation, we use a database of clinical images and the ground truth annotated by experienced human observers with support of a semi-automatic software.

3.5.1 Dataset

Though this method is designed to segment empty catheter in fluoroscopic images, we evaluate our empty catheter segmentation algorithm using a dataset fluoroscopic images and record images to determine its robustness to image quality. The dataset consists of 1279 fluoroscopic images belonging to 10 sequences taken from examinations of 6 patients. While it consists of 630 record images belonging to 7 sequences from 5 patients. These images were acquired at frame rate of 15 fps. Considered images of angioplasty exams depict large variability because of patients' body mass index (BMI), noise levels, different anatomical backgrounds, occasionally presence of pacing leads, stents, staples, sternal wires (see Table 4.2).

3.5.2 Ground truth construction

Catheter appears as low contrasted tubular structures in X-ray images. We decided to have the centerline of the catheter as a reference point to evaluate the performance of our segmentation. An internally developed (image similarity based) semi-automatic software for curve tracking is used by human operators to mark, track and correct the centerline, forming a curve which forms the ground truth for empty catheters in fluoro images along the temporal sequence. For each frame in a sequence, a centerline is a curve in 2D image space, which is then sampled in a series of equidistant pixels as $\mathcal{C}_{gt} = (g_1, \dots, g_m)$.

3.5.3 Segmentation

The automatically detected empty catheter is a cluster of curve blobs. As described in previous section, a polynomial curve is fitted to these blobs which is then sampled in a series of equidistant pixels given by $\mathcal{C}_{seg} = (s_1, \dots, s_n)$. Right column in Figure 3.10a shows the estimated centerline. In \mathcal{C}_{gt} and \mathcal{C}_{seg} , the sampling distance between two consecutive points is 1 pixel (0.2mm).

3.5.4 Evaluation measures

In this work, we want to evaluate our ability to locate the empty catheter. We quantify the proximity between the two objects: the reference centerline, curve marked as ground truth, and the estimated centerline, fitted polynomial curve. This metric of proximity is then analyzed using the precision and recall formalism. Precision is defined as fraction of correctly detected catheter. As explained in Figure 3.9a, the matched detection is denoted as true positive, emphasizing the fact that the segmentation algorithm has indeed found the catheter. The unmatched detection is denoted as false positive, because the detected catheter hypotheses are incorrect. Similarly, recall is fraction of reference centerline (ground truth) which is retrieved and retained in the estimated centerline. Figure 3.9b shows the matched reference (true positive) which is correctly retrieved ground truth points. Such centerline line based evaluation methods are employed for evaluation of road extraction algorithms in photogrammetry and remote sensing [Wiedemann et al., 1998].

Precisely explaining our implementation, for each image, we quantify the proximity between: the series $\mathcal{C}_{gt} = (g_1, \dots, g_m)$ of ground truth points and the series $\mathcal{C}_{seg} = (s_1, \dots, s_n)$ of the points extracted from segmentation. To this end, we consider the minimal distance from a point x to a series of points $\mathcal{C} = (c_1, \dots, c_\ell)$ as,

$$\delta(x, \mathcal{C}) = \min\{d(x, c_i) \mid i \in \{1, \dots, \ell\}\} , \quad (3.11)$$

where d is Euclidean distance. Based on this measure, a point s_i of the segmented catheter curve \mathcal{C}_{seg} is considered as correctly classified (true positive) when,

$$\delta(s_i, \mathcal{C}_{gt}) \leq \eta . \quad (3.12)$$

Each point s_i which does not satisfy this equation, is classified as false positive.

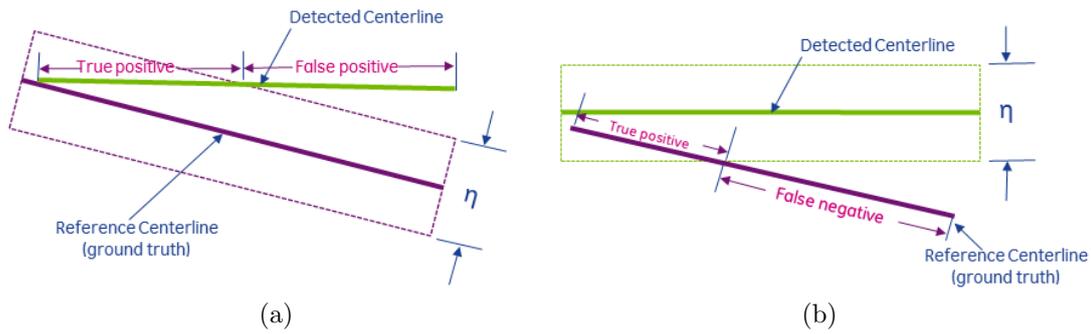


Figure 3.9: Evaluation principle to compute precision and recall. (a) Matched detection (extracted centerline); (b) Matched reference (ground truth)

These sets of true and false positive points are used to compute precision (Figure 3.9a). In this regards, we enumerate the number of true positives ($\#TP$) and false positives ($\#FP$) in \mathcal{C}_{seg} . Thus we compute the *Precision* as the fraction of segmented points correctly classified,

$$\text{Precision} = \frac{\#TP}{\#TP + \#FP} \quad (3.13)$$

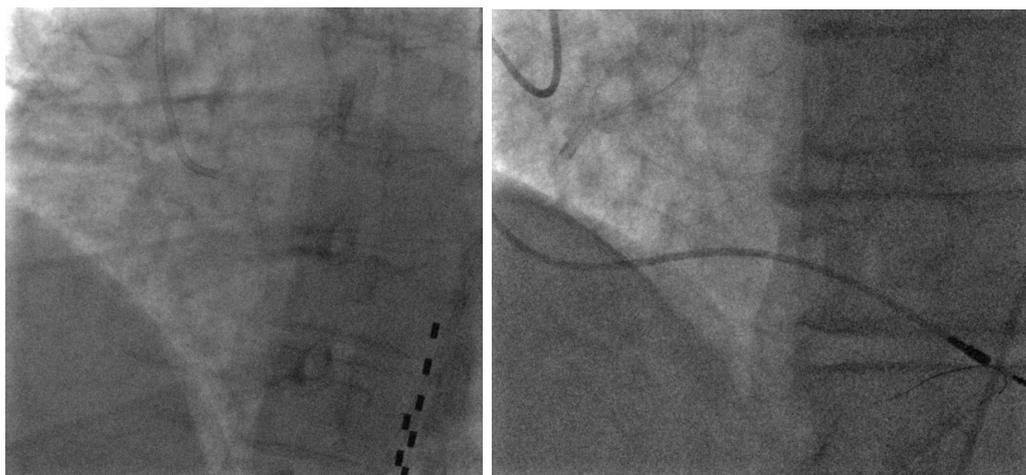
Similarly, a point g_i of the ground truth \mathcal{C}_{gt} is considered as correctly retrieved when,

$$\delta(g_i, \mathcal{C}_{seg}) \leq \eta . \quad (3.14)$$

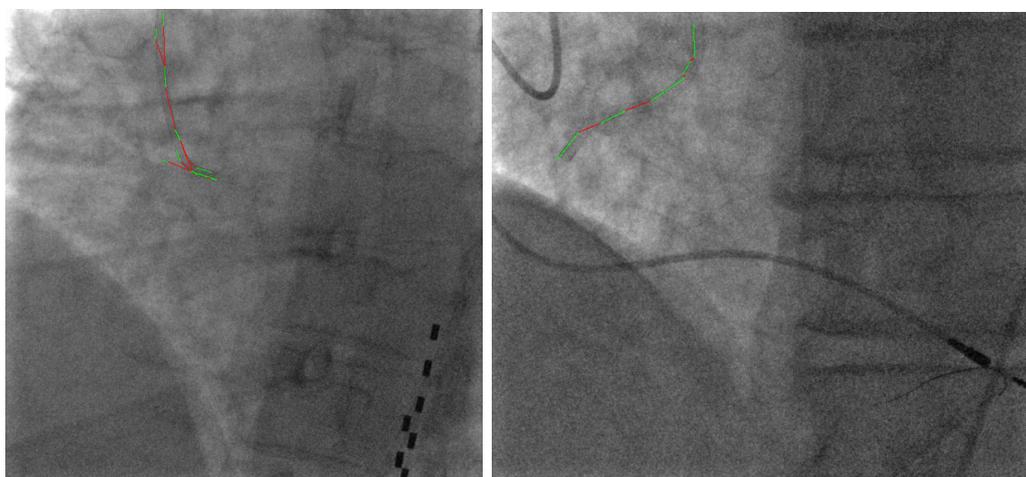
On this basis, each point $g_i \in \mathcal{C}_{gt}$ is classified as true positive or false negative to calculate the number of true positives ($\#TP$) and false negatives ($\#FN$) (Figure 3.9b). From this information, we compute *Recall* as the fraction of ground truth points correctly retrieved,

$$\text{Recall} = \frac{\#TP}{\#TP + \#FN} \quad (3.15)$$

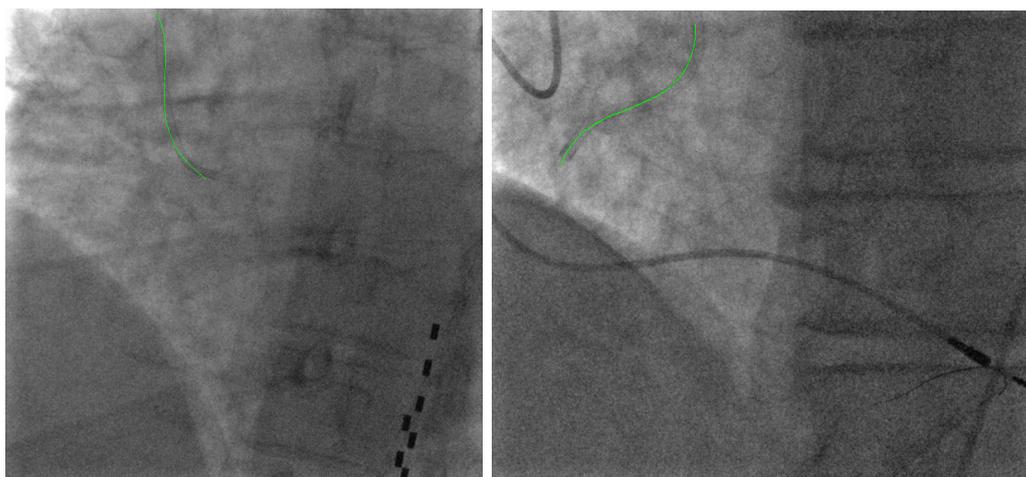
The value of η used for evaluation is based on the the standard diameter of an empty catheter in the image plane (here $\eta = 24$ pixels (4.8mm)). We also analyze the precision and recall for various possible values of η .



(a) Input images



(b) Selected clusters of curve blobs



(c) Estimated centerline of catheters

Figure 3.10: Results of empty catheter segmentation in X-ray fluoroscopic images. (a) input fluoroscopic images from two patients, (b) selected cluster of curve blobs from hierarchy (c) fitted curve on the cluster (estimated centerline).

3.6 Results and discussion

We evaluate our empty catheter segmentation algorithm using a dataset of 1279 fluoroscopic images and 630 record images of cardiac angiographic sequences of patients. The segmentation of empty catheters in fluoroscopic images with different image qualities and different anatomical and interventional contents are shown in the Figure 3.10. Figure 3.10a shows two fluoroscopic images which are input of the segmentation algorithm, and in Figure 3.10b, the corresponding input images are overlaid with the selected cluster of curve blobs. Whereas, Figure 3.10c shows the fitted curve for the selected cluster of blobs, this curve is considered as an estimation of the centerline of the catheter. Figure 3.10c portrays empty catheter segmentation in presence of other elongated interventional and anatomical objects.

Figure 3.12a depicts the results of empty catheter segmentation in fluoroscopic images from three different patients. These images illustrate the potential of the segmentation method, where empty catheter is detected in spite of the presence of other elongated objects like pacing leads, sternal wires. These fluoroscopic images also have disturbing anatomical contents like the spine. Indeed, some internal structure of the vertebra bodies may take part of the appearance of catheter because of contrast and curvilinear outlook. Figure 3.12b shows the result of our segmentation method on the record images from three patients. Rightmost image in this figure refers to the point in the sequence, towards the end of an injection in the cardiac artery and the fading injection is still visible. Whereas, left and middle images in Figure 3.12b belong to rotational record sequence and contains sternal wires.

For fluoroscopic images, we assess two versions of our automatic algorithm with and without temporal feedback using the defined evaluation measure. According to the evaluation measure defined in Section 3.5.4, precision and recall can be computed for obtained centerline of each image. In order to analyze the performance for a sequence, we report the mean precision and mean recall over all the images in the sequence. Similarly, we report weighted mean precision and recall for the complete dataset where the weights are the number of images in each sequence. In 1279 fluoroscopic images, the weighted mean precision and recall without temporal feedback are $62.40 \pm 18.86\%$ and $55.84 \pm 18.2\%$ respectively. With temporal feedback, this weighted mean precision and recall improves to $83.85 \pm 23.29\%$ and $67.87 \pm 19.53\%$ respectively. The weighted standard deviation of precision and recall with temporal feedback is more because per sequence performance improves more for sequences with higher rates of precision and recall without temporal feedback. For record images, we analyze the performance with temporal feedback. The segmentation is better due to better image quality as compared to fluoroscopic images. With temporal feedback, the weighted mean precision and recall values are $88.2 \pm 13.54\%$ and $71.5 \pm 12.78\%$. Some record sequences were obtained during C-arm spin, leading to some artifacts due to occlusion with sternal wires.

In case of fluoroscopic images, we graphically analyze the precision and recall over complete dataset by drawing the boxplots (3.11) for precision and recall at different thresholds (η in Equations 3.12 and 3.14). In these graphs, the red line connects the medians at various thresholds. The weighted mean values at each

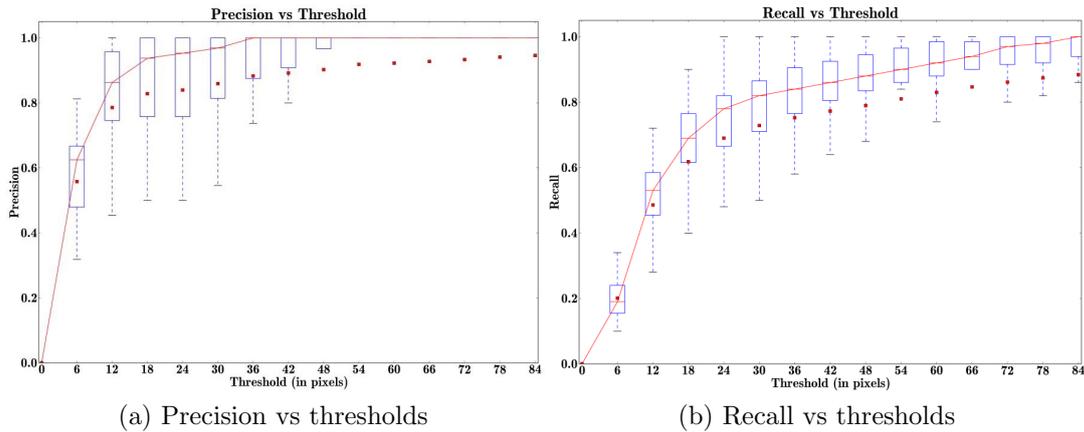


Figure 3.11: Box plots of precision and recall at different thresholds in fluoroscopic images.

threshold is marked by red square in the corresponding box plot. The whiskers in these box-plots refer to $1.5 \times (\text{Quartile Range})$ value. Among these informative plots, we are particularly interested in the precision and recall at threshold of 24–30 pixels, which is slightly more than the width of the catheter which varies between 20 – 24 pixels. In these two plots, it can be seen that the mean value is lower than the median value. This is mainly because of the poor performance in some patient sequences namely B2, D1, and E2.

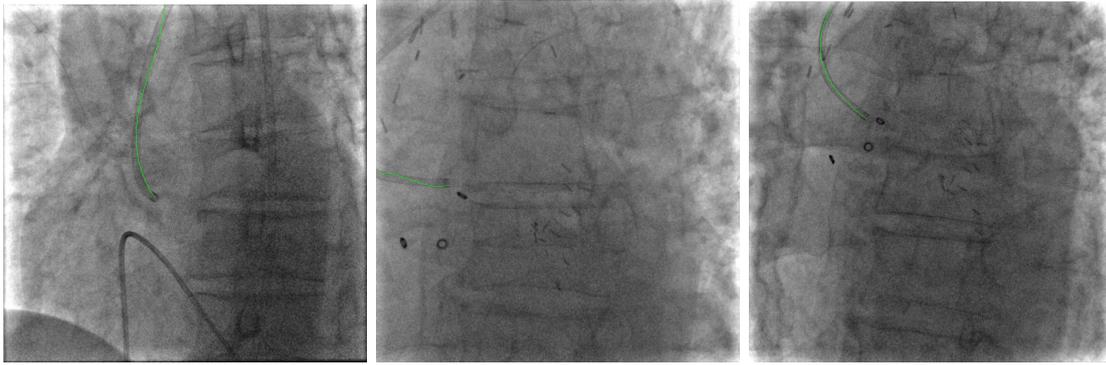
In a detailed per sequence analysis (Table 3.1), we notice that in few sequences the performance of our algorithm is hampered because of some factors such as the patient’s body mass index (BMI), catheter appearing over the spine making it less visible, several sections of catheter in the field of view (FOV) leading to multiple apparent catheters (*e.g.* left and middle images in Figure 3.12a). Low precision and recall was observed in sequences B2, D1 and E2 (in Table 3.1) due to multiple section of the same catheter in the FOV and high patient body mass index. In sequence B2, our proposed algorithm fails to identify the desired section of catheter with the tip, among the two sections. Figure 3.13a depicts a frame from sequence B2, where ground truth is marked in green and detected catheter is marked in red. However, in sequence E1 the undesired section of catheter (without the tip) was above the spine making it difficult to be detected. Hence, the precision and recall for this sequence is 96.54% and 71.10% respectively, leading to successful segmentation of desired section of catheter with tip.

In sequence D1, the algorithm fails due to high patient BMI of 39.7, resulting in high noise level and very low contrasted catheter. It can be seen in Figure 3.13b, that the catheter is barely visible to naked eyes. We also noticed that position and orientation of the gantry affects the image quality and the performance of our algorithm. In further experiments we observe that the precision and recall rate is stable when parameters of the edge weights are changed in a range of $\pm 20\%$, which is encouraging regarding the robustness of the approach which is not dependent of very precise parameter setting. A small set of 30 images from 4 sequences (A1, B1,

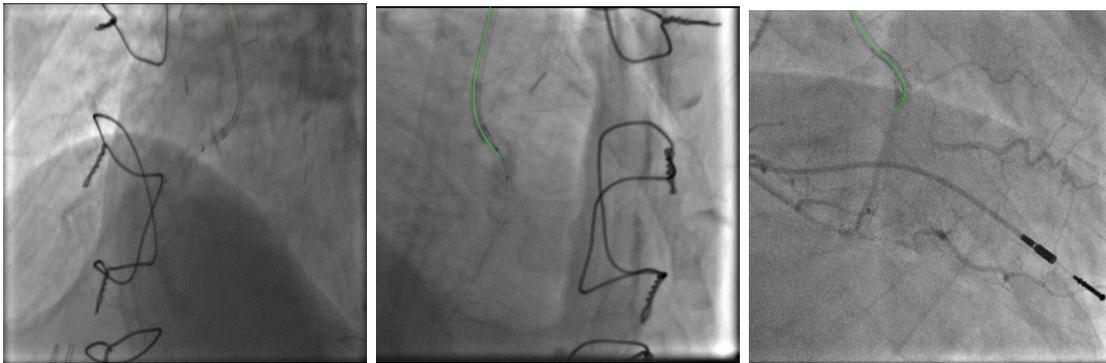
Table 3.1: Patient wise analysis of sources of variance of the performance of empty catheter segmentation algorithm

Patient	Sequence	Number of images	Body Mass Index (BMI)	Average Noise Level(*)	Catheter over spine	Several sections of Catheter in FOV	Without temporal feedback		With temporal feedback	
							Precision	Recall	Precision	Recall
A	A1	115	30.1	23.18	No	No	66.11%	66.96%	89.42%	74.35%
	B1	178	29.7	20.64	No	No	75.37%	72.85%	92.26%	87.4 %
	B2	30	29.7	20.44	No	Yes	6.24 %	9.45 %	8.43 %	13.98%
C	C1	118	35.5	19.80	No	No	52.16%	35.7 %	75.88%	52.25%
	C2	218	35.5	20.82	No	No	58.09%	51.33%	89.32%	72.31%
	C3	159	35.5	20.43	No	No	72.57%	62.94%	90.35%	72.6 %
D	D1	59	39.7	27.15	Yes	Yes	8.34 %	7.21 %	11.14%	11.24%
	E1	290	22.0	26.09	No	Yes	73.02%	65.53%	96.86%	74.01%
E	E2	42	22.0	24.77	No	Yes	34.55%	32.9 %	48.78%	36.77%
	F1	70	24.4	17.34	No	No	73.39%	60.74%	95.78%	73.47%
Mean values							62.40%	55.84%	83.85%	67.87%

(*) Image dynamic is 256 levels



(a) Fluoroscopic images



(b) Record images

Figure 3.12: Results of empty catheter segmentation in X-ray images. Detected centerlines in (a) fluoroscopic images of three patients, (b) record images of three other patients.

B2, C1) was used for tuning parameters in the full algorithm development. Once the development was completed, we built a large database of images with ground truth. Optimizing the α and β parameters of the blob pair weight function (refer equation 3.7) on 650 images (instead of 30) from 4 sequences (A1, B2, C1 and D1) slightly improves the results (3.45% Recall / 6.20% Precision). Our evaluation measure and the ground truth are used for this optimization step.

The average execution time per image is 0.58 seconds on a Intel® Core™ i7 – 4810MQ CPU. The software has good potential for further optimization.

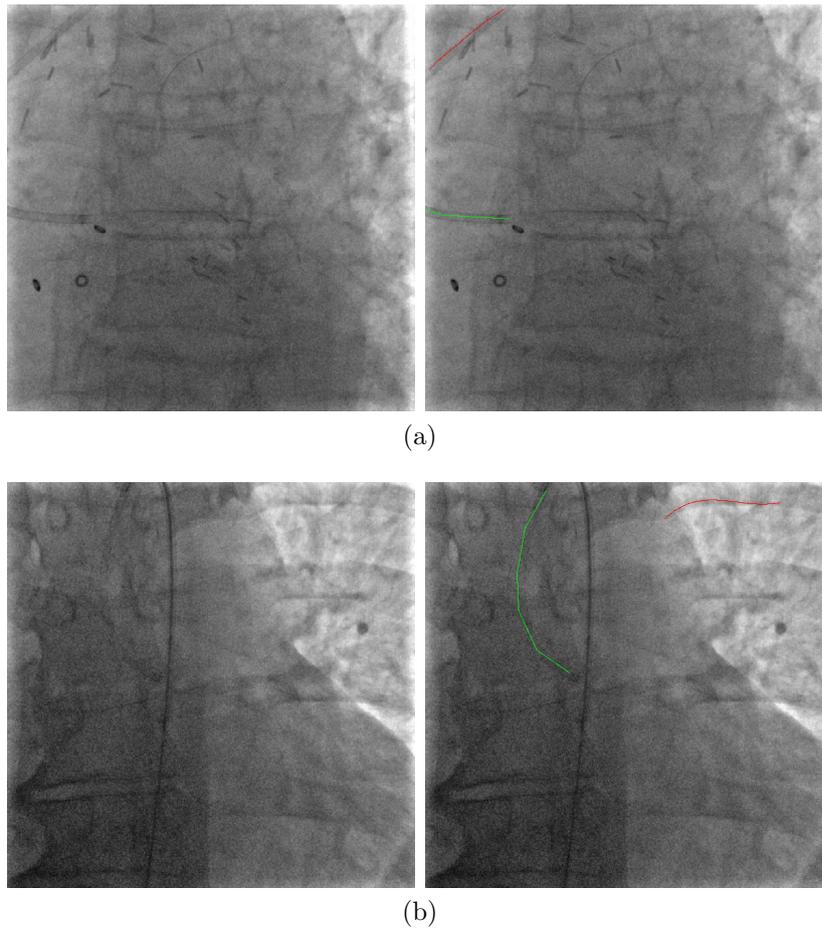


Figure 3.13: Failed segmentation cases. In (a) and (b), left image shows the input image and right image shows the ground truth in green and the falsely detected catheter in red. See Section 3.6 for details.

3.7 Conclusion of catheter segmentation

In this chapter, we studied the challenging problem of detecting and locating the empty catheter in fluoroscopic images. We also presented the results of segmentation in record images. Empty catheter segmentation is a very challenging problem taking into consideration the appearance of the catheter in fluoroscopic images. To achieve our goal, we developed a novel structural scale-space in the form of a hierarchy of deep connected components, one of them being selected as empty catheter. Our experimental results are very encouraging, showing that it is indeed possible to locate with good precision the empty catheter in such noisy images. These results also open the doors for PCI procedure modeling since empty catheter is an important landmark in these images. Indeed, using a similar strategy in the scale space framework, we aim to simultaneously detect other landmarks, such as guide wire tip, marker balls, or balloons that are well contrasted. Therefore, our next chapter includes segmentation of these objects by handling in a common scale-space framework. These segmentations can contribute to the PCI procedure modeling.

VOIDD: vessel of intervention dynamic detection and tracking

Contents

4.1	Introduction	76
4.1.1	Clinical interest	76
4.1.2	Related work	77
4.1.3	Major contributions	80
4.2	Data and preprocessing	81
4.2.1	Reference sequence	81
4.2.2	Fluoroscopic image stream	82
4.2.3	Iso-phase image pairing	82
4.2.4	Table panning and gantry rotation compensation	84
4.3	VOIDD algorithm	85
4.3.1	General tracking framework	87
4.3.2	Feature pairs extraction	87
4.3.2.1	Vessel centerline extraction	87
4.3.2.2	Tip candidate extraction	89
4.3.2.3	Matching	92
4.3.3	Track assignment cost	93
4.3.4	Vessel of intervention refinement	96
4.4	Dataset and evaluation measures	99
4.4.1	Dataset and expert annotations	99
4.4.2	Tip detection evaluation measure	101
4.4.3	Track evaluation measure	101
4.4.4	VOI evaluation measure	101
4.5	Results	102
4.6	Conclusion and future work	109

In the chapter, we present a fully automatic tracking algorithm, called VOIDD, to detect the vessel of intervention (VOI), which is treated during the procedure, by combining information from a vessel image with contrast agent injection and fluoroscopic images acquired during guidewire tip navigation where the VOI is not visible. The crux of the algorithm lies in creating and tracking feature pairs that are spatially and temporally consistent. In order to form such feature pairs, we also present

1. a robust guidewire tip segmentation method,
2. a vessel tree extraction algorithm, and
3. a matching procedure allowing to make smart correspondences between detected guidewire tips and vasculature.

Such capabilities can be used for varied applications like automatic documentation of PCI procedures, automatize preparation of enhanced images and also to optimize the image acquisition to reduce the use of ionizing radiation or amount of contrast media. We provide evaluation methodologies to characterize the correctness of the guidewire tip detection and correct identification of the VOI. The method is evaluated on a 15 clinical sequences dataset consisting of 9989 fluoroscopic images and 150 VOIs annotated by a trained observer. Our results show that VOI can be robustly obtained from X-ray images acquired during the PCI procedures. We also show that our method is able to detect the sequences of guidewire navigation, that is the appearance of guidewire tip in field of view, with high accuracy and we assess its robustness with respect to different patient and imaging conditions.

4.1 Introduction

4.1.1 Clinical interest

As mentioned in Chapter 3, the catheter is placed at the ostium of each of the left and right coronary tree, one after another, thus preparing for the vessel diagnosis phase. The vessel diagnosis is one of the first phases in the PCI procedure. It includes the visualization of the lumen of the coronary arteries which is achieved by injecting an iodine based contrast media at the ostium of each coronary tree (left and/or right). The contrast media flows through the catheter into the coronary vessels and is rapidly washed out by the blood flow. The propagation of the contrast media is documented under X-ray radiation to acquire the angiographic sequences which are further used to diagnose the stenoses. The procedure is overall beneficial to the patient but has several side effects that we aim to reduce. However, the tolerance to the contrast agent is limited to some amount. The other side effect is the use of ionizing radiation which affects both the patient and the medical team present in the interventional room including the interventional cardiologist.

Vessel of intervention: The *Vessel of intervention* (VOI) is a branch of the coronary vessel tree between the ostia and the distal end of the vessel across the

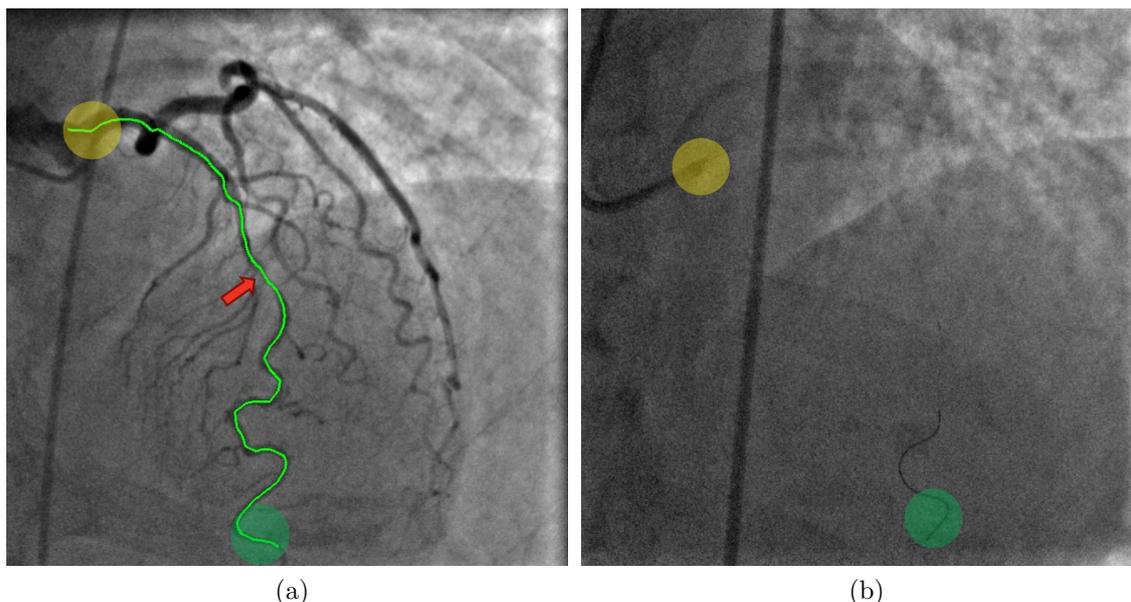


Figure 4.1: Vessel of intervention definition: a branch of coronary vessel tree in which the guidewire is navigated and is between ostia (yellow region) and the distal end of the vessel (green region) across the coronary lesion (red arrow). (a) shows the VOI in injected vessel image and (a) shows the guidewire tip at the distal end of vessel.

coronary lesion. Coronary lesions are treated by navigating a guidewire through the VOI (during the *guidewire navigation phase*), followed by implantation of a stent at the lesion in the VOI. Figure 4.1 illustrates this vessel of intervention in injected vessel image and guidewire navigation image. In the work presented here, we develop methods for automatic detection of VOI by combining the information from X-ray image sequences acquired at different steps of the procedure. Our aim is to automatically identify the guidewire arrival at the ostia of the coronary vessel tree and to determine the VOI which is going to be treated in the following steps of the PCI procedure, such as lesion reparation with angioplasty balloon, stenting, post-dilatation. The detected vessel of intervention can be used as a region of interest to segment forthcoming objects like marker balls or balloon. One of the goals of PCI procedure modeling is the detection of the procedure phases. We associate the detection of the guidewire navigation phase to the task of identification of the arrival of guidewire tip in the field of view. Thus, giving the capability to the imaging equipment to monitor precisely the phases of the procedure can lead to improvement of the overall workflow of these procedures. In the long term, this capability could be used to optimize the image acquisition to reduce the amount of dose or contrast media employed during the procedures.

4.1.2 Related work

Keeping in mind the goal of PCI procedure modeling, detecting the vessel of intervention can lead to significant semantic information about the potential region

of interest for the following part of the procedure. However, there has not been any work in literature with a goal similar to this one. Some of the works with similar clinical interests are some applications from the manufacturers of interventional suites. Some manufacturers of interventional suites propose applications which facilitate the visual appreciation of the relationship between the guidewire and the vessel. The main idea is to combine a sequence of consecutive injected images, which visualize the vessel along a cardiac cycle, with the images obtained during tool navigation as briefed by [Dannenberg et al., 2016]. The images at these different times are paired mostly based on the ECG. Breathing motion may be compensated by comparing moving structures in the images [Dannenberg et al., 2016]. The slight deformation of the arteries caused by the introduction of the guidewire is not compensated. The detailed methodology and quantitative measurements on the obtained accuracy have not been reported so far for this challenging problem. Though these methods maybe useful for navigation of the tool, but they may not be easily extended to detect the region of interest (here, VOI) in the injected image.

From a more technical standpoint, looking at the task of identification of the vessel of intervention in angiographic images, significant work exists in literature with a different aim and a workflow than our purpose. Most of such works have been done to co-register between different modalities, such as X-ray angiographic sequences with

1. intra-vascular ultrasound (IVUS) or with
2. optical coherence tomography (OCT)

which are involved in some of the PCI procedures. The co-registration between angiographic sequences and IVUS has been studied by [Wang et al., 2013; Prasad et al., 2016; Frimerman et al., 2016]. In these methods, the position of IVUS transducer during pullback is mapped to segmented vessel branch in ECG triggered angiographic image. The outcome of these methods is registration of corresponding point in the IVUS plane to a point along the segmented vessel branch. The clinical studies and methods for co-registration of coronary angiographic sequences and OCT has been presented by [Tu et al., 2011; Hebsgaard et al., 2015; Koyama et al., 2015]. These co-registration methods provides the interventional cardiologist with detailed information about vessel size and plaque size at every position along the vessel of interest, making themselves suitable tools during the actual intervention. [Prasad et al., 2016] have shown that it is possible to search the correspondence between a location identified in the low-dose X-ray (fluoroscopic) images acquired during tool navigation and the high dose X-ray (cine) images which depict the injected vessels. The addressed clinical need is the registration of intra-vascular images acquired with a sensor placed along the guidewire with the vessel. By this means, the operator can easily correlate the readings of the angiographic images and the intravascular images signals. In this situation, a full application is developed with a specific acquisition workflow with the different steps of the image acquisition and processing being done based on landmark points and appropriate images selected by an operator. One of the important step in these methods involve manual

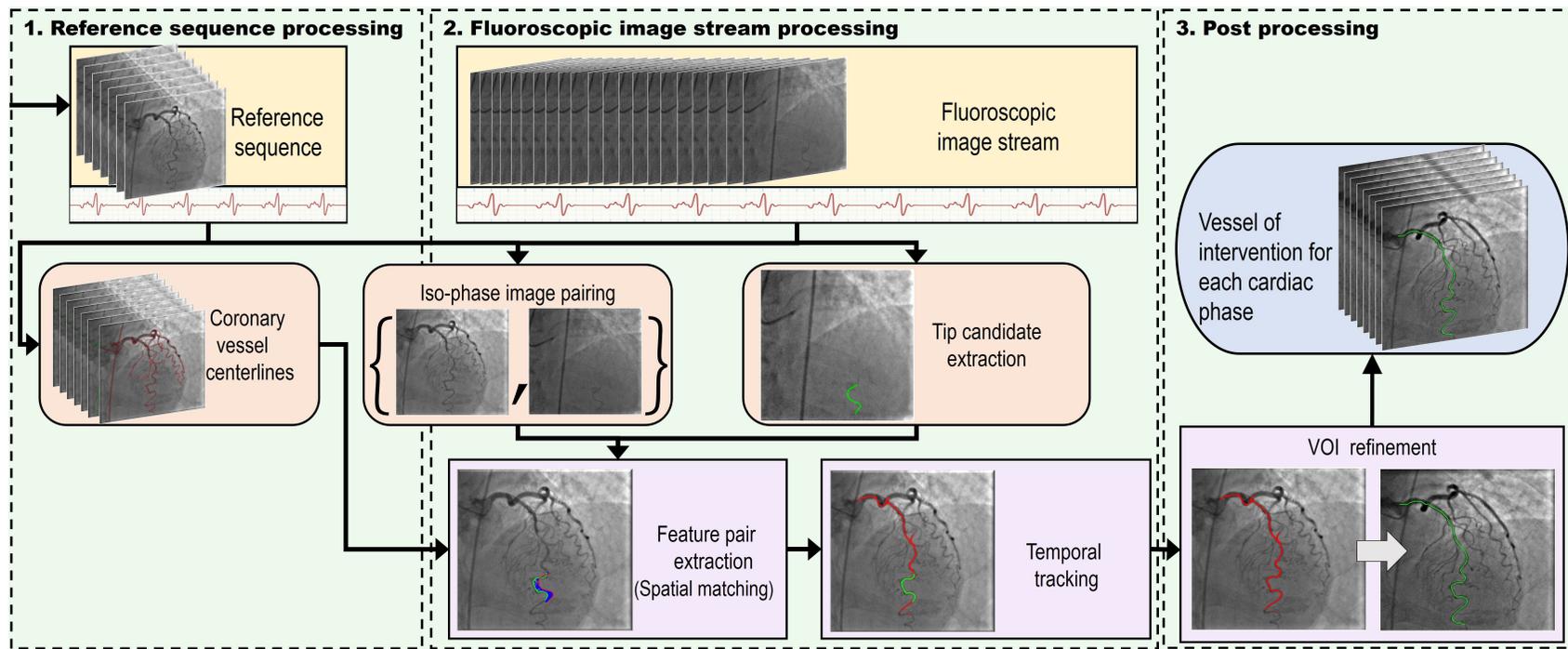


Figure 4.2: Vessel of intervention detection system workflow

identification within the diagnostic angiogram of the vessel that is subject to IVUS examination.

Our tracking framework aims at determining the corresponding position of the guide-wire tip in the injected vessels. The significant prior information required for our tracking framework includes guidewire tip segmentation, coronary vessel centerlines and the matching between segmented guidewire tips and centerlines of the coronary vessels. Several authors have worked on the task of segmenting the guidewire [Vandini et al., 2017; Chen et al., 2016; Wang et al., 2017, 2009]. For PCI application as in [Honnorat et al., 2012], the weak contrast of the guidewire body makes the task very challenging. Therefore, we decide to use the tip of the guidewire to extract information regarding the advancement of the guidewire in the vasculature. In order to extract coronary vessel centerline, Hessian based vessel enhancement techniques [Frangi et al., 1998; Kirbas and Quek, 2004; Zhang, 2010] have been largely used in the literature. These techniques are based on the basic assumption that the vessels are locally linear tubular structures. They are often composed of two steps: 1) enhancing the vessels in the images, 2) segmenting vessel centerlines.

4.1.3 Major contributions

The main contribution of this chapter is the proposition and the assessment of a method, called VOIDD, to automatically detect the so-called vessel of intervention, from the images acquired during the initial phases of the PCI procedure. We combine information at these initial phases from two different X-ray image sequences: i) the cine images from *reference sequence*, a subset of angiographic sequence, which are injected with contrast agent to depict the vasculature during complete cardiac cycle and ii) the fluoroscopic images from *fluoroscopic image stream*, which are acquired following the reference sequence to aid navigation of various tools and especially the guidewire. These phases of the procedure are introduced in the 1.2.2.2 and in Figure 1.7. These X-ray image sequences are further detailed in Section 4.2 to lay the basis for this Chapter. The different processing steps of the method are shown in Figure 4.2. More precisely, this algorithm is able to recognize from the fluoroscopic image stream, the period corresponding to the guidewire navigation and to exploit it to determine the vessel of intervention location (see Fig. 4.3) without adding any constraint to the procedure workflow. In order to reach this goal, a general tracking algorithm is proposed and explained in Section 4.3. This algorithm relies on features extracted from the two considered types of X-ray image sequences. These features consist of coronary vessel centerlines extracted from reference sequences and of guidewire tip location candidates detected in fluoroscopic image streams. Hessian based vesselness techniques are used to obtain the vasculature from reference sequences. Guidewire tip detection is tackled with advanced approaches involving the use of min tree [Salembier and Wilkinson, 2009]. Fréchet distance based curve matching approaches derived from [Benseghir et al., 2015] are used to match the guidewire tip with detected vessels. In Section 4.4, we present the evaluation methodologies designed to characterize the correctness of the guidewire

tip detection and the correct identification of the VOI location. These developments are assessed on 15 clinical sequences dataset from 14 patients and comprising 9989 images with expert annotations. The results are presented and discussed in Section 4.5.

4.2 Data and preprocessing

This section describes the data used in the different stages of the algorithm. This section also details the pre-processing and synchronization of image data with respect to non-image data. The system workflow, as described in Figure 4.2, comprises of two main stages: the reference sequence processing and the fluoroscopic image stream processing followed by a post-processing stage. The reference sequence, consisting of images of contrast-enhanced vessels, is described in Section 4.2.1. The fluoroscopic image stream for navigation of tools (here, guidewire) is described in Section 4.2.2. Section 4.2.3 explains the synchronization of data from different sequences performed to pair images according to same cardiac phase. Section 4.2.4 describes pre-processing of images with homography transformation to align images in the same reference plane.

4.2.1 Reference sequence

The purpose of the angiographic sequence is to document contrast enhanced cardiac vasculature in a X-ray image sequence using the cine mode of the interventional system. In cine mode, the system is set to deliver images with a quality sufficient to support the operator in his assessment of the vasculature. In general, angiographic sequence may have around 10-20 images with contrast agent highlighting the complete vasculature. The number of images with well contrasted vasculature depends on the amount of contrast agent injected. Standard algorithm as one proposed by [Hoffmann et al., 2015] can then be used to identify such subset of the images where the coronary images are well opacified with the contrast agent. In this subset, a *reference sequence* \mathcal{R} of reference images is then selected that covers a full cardiac cycle. The ECG acquisition, as mentioned in Section 1.2.1.2, maps the X-ray images to the ECG data. This mapping aids the selection of the reference sequence covering full cardiac cycle. In the interventional systems considered for this work, the acquisition frame rate of cine images is 15 frames/second. The normal heart rate being 75-80 beats per minute, a full cardiac cycle takes 0.7-0.8 seconds of time. Thus, the reference sequence is selected to include best opacified 10-15 images (based on acquisition frame rate and heart rate). In the following, any image of this reference sequence is called a *reference image*, and is denoted by R . In relation to the system workflow, the reference sequence can be processed at the acquisition of the full angiographic sequence.

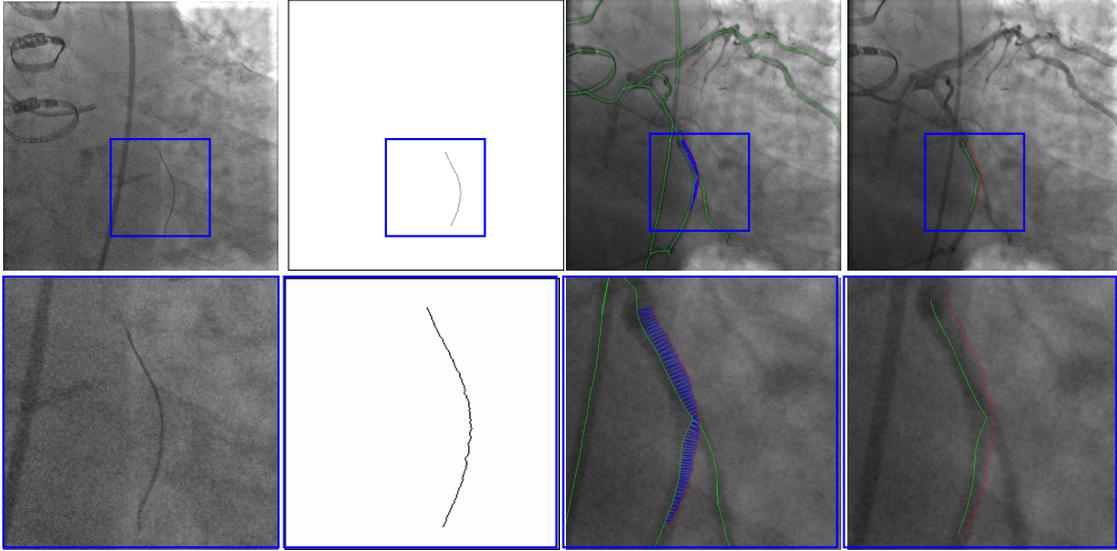


Figure 4.3: VOIDD: (from left to right) Input image f ; centerline of segmented guidewire tip; tip candidate (red) matched to vessel centerline (green) marked by pairings (blue); corresponding location (green) of guidewire tip (red) inside vessel.

4.2.2 Fluoroscopic image stream

The reference sequence acquisition is usually followed by the navigation of guidewires and other tools through the catheter until the distal end of the VOI. A guidewire, a very thin (wire-like) object, usually has a stiff body and a floppy tip. The tip being the distal section of 20 mm length, aids the navigation of guidewire. It is enough radio opaque to be seen with fluoroscopy, low-dose X-ray mode (input image in Figure 4.3). The acquisition rate of fluoroscopy mode of interventional system (used in this study) is also 15 fps. The *fluoroscopic image stream*, denoted by \mathcal{F} , contains the fluoroscopic images acquired during navigation of the guidewire, which is navigated from the ostia of the coronary vessels down to the distal part after crossing the lesion (as shown in Figure 4.1b). The fluoroscopic image stream is most of the times split in several sequences corresponding to stepping on/off the X-ray pedal by the physician. Some of these sequences contains images where guidewire tip is not visible in the field of view as it has not yet arrived at the ostia. In the following, any image of such fluoroscopic image stream \mathcal{F} is called a *fluoroscopic image* and is denoted by F . The ECG signal is recorded in parallel to these images as well. Each individual image of the fluoroscopic image stream \mathcal{F} shall be paired to an image of the reference sequence \mathcal{R} . It is mostly done based on ECG signal as explained in Section 4.2.3. Unlike the processing of the reference sequence, the processing of fluoroscopic image stream shall be casual.

4.2.3 Iso-phase image pairing

The cardiac motion causes periodic position and shape changes in the coronary vessels and these deformations apply to the guidewire tip which is navigated in the

lumen of these vessels. Therefore, the appearance of guidewire tip in the 2D projected X-ray image is continuously changing. Consequently, we elected to pair each individual image of the fluoroscopic image stream with an image from the reference sequence, where the depicted vasculature and guidewires have corresponding deformations induced by cardiac motion. In order to perform this pairing, we use the patient ECG as a mean to identify for each image a phase in the cardiac cycle and pair the images based on this phase. The ECG is recorded with the X-ray images and is analyzed to identify the QRS complex which marks the start of a cardiac cycle. These cardiac phases are indicated as the percentage interval between two R-waves (R-R interval) from the QRS complex. In this work, a cardiac cycle is divided into Φ cardiac phases ($\Phi = 10$ here). Therefore, we select Φ images from the reference sequence which correspond to the Φ divisions of the cardiac cycle. This set of images is given by the sequence,

$$\mathcal{R} = \{R_\phi \mid \phi = \{1, \dots, \Phi\}\} . \quad (4.1)$$

The cardiac phase of each image F of the fluoroscopic image stream is also determined by analyzing the recorded ECG. It is then paired to the retained image of the reference sequence which has the closest phase. We refer such pair of images F and R at cardiac phase (ϕ), as an *iso-phase image pair*, denoted by (F, R) . The set of all iso-phase image pairs for given \mathcal{F} and \mathcal{R} is

$$\{(F, R) \mid F \in \mathcal{F}, R \in \mathcal{R}, \phi(R) = \phi(F)\} , \quad (4.2)$$

where $\phi(I)$ is the cardiac phase of the image I .

Despite of iso-phase image pairing, there may still exist some spatial inconsistencies between the two images in an (F, R) because of numerous reasons as listed below:

- breathing motion, as the images in the iso-phase image pair may come from different breathing phase;
- vessel deformations, due to the presence of stiff guidewire body in the vessel
- tablepanning and gantry rotation, as the physician may change the settings of the interventional systems between the acquisition of the reference sequence and guidewire navigation phase.
- ECG phase mismatch, due to electronic delay of the coupled ECG machine.

The inconsistencies induced by the tablepanning and gantry rotation can be compensated with geometrical transformations. However, to cope up with the inconsistencies induced by remaining factors, it is necessary to design a matching and tracking algorithm to coherently track the guidewire tip in the reference image by taking into account all these factors. In order to curtail the effect of tablepanning and gantry rotation, we pre-process images from reference sequence and fluoroscopic image stream with table panning compensation, as explained in next section.

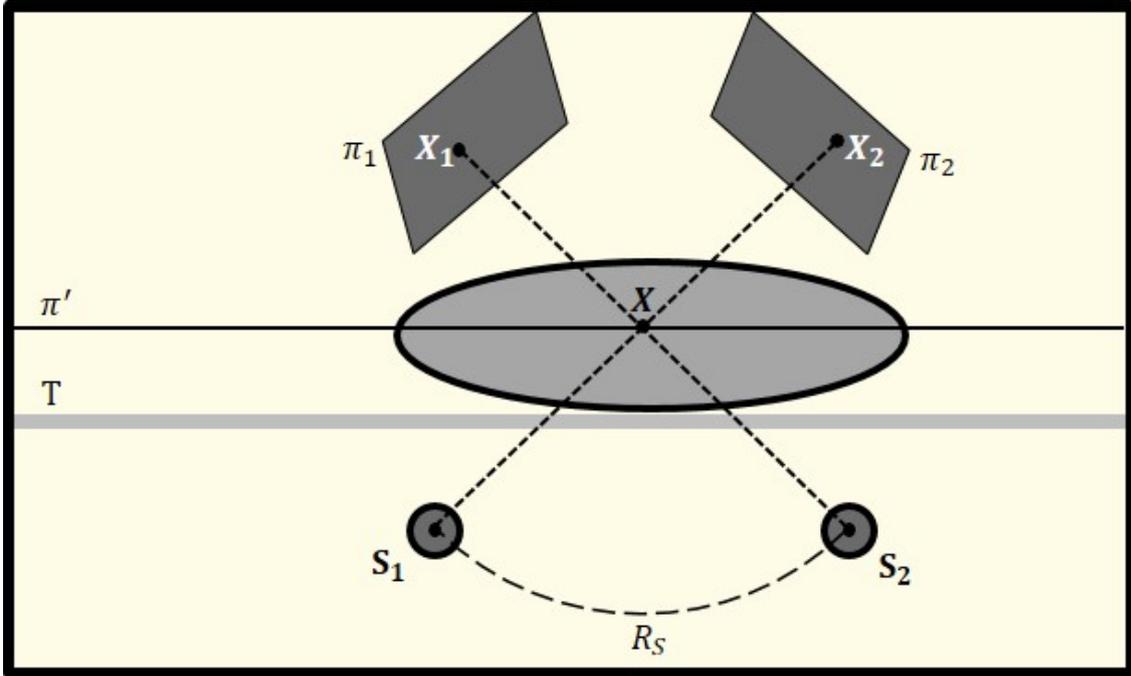


Figure 4.4: Projective geometry in case of gantry rotation: X_1 and X_2 are projections of a 3D point X in image (detector) planes π_1 and π_2 respectively. S_1 and S_2 are the positions of the X-ray source in these two cases.

4.2.4 Table panning and gantry rotation compensation

During the navigation of the tools, the physicians commonly maintain the settings of the interventional systems such as gantry angulation and the table position to be the same as or close to the ones of the reference sequence. However, in some cases the table position and the gantry angulation are not the same for reference sequence and the fluoroscopic image stream. As a consequence, the appearance of the anatomy between the two sequences is changed. We figured out that the knowledge of the system configuration enables to partially correct this difference of appearance. In this section, we explain the projective transformation.

A camera model is used for perspective projection to describe how the observed 3D point project onto an image. Let $(X, Y, Z, 1)^t$ be a 3D point in homogeneous coordinates and let $(x, y)^t$ be the image of this point in the camera, then the camera model is a linear map between the homogeneous point coordinates, given by

$$\begin{pmatrix} sx \\ sy \\ s \end{pmatrix} = P \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}, \quad (4.3)$$

where P is a 3×4 projection matrix (or projective operator) with 11 degrees of freedom and s is any scalar. Assuming now that we consider only the 3D points belonging to a given plane. As a consequence, X, Y, Z can be expressed linearly

from R, S , where (R, S) are the coordinates of a point defined in a referential attached to this plane. So $(X, Y, Z, 1) = G(R, S, 1)$ where G is a 4×3 matrix with its last line being $(0 \ 0 \ 1)$. By introducing these elements in Equation 4.3, it can be further simplified to $(x \ y \ 1) = H(R, S, 1)$. H is a 3×3 homogeneous matrix.

Figure 4.4 exemplifies the projective geometry during X-ray acquisition. Let us consider two acquisitions done in different geometry and one plane. Here, S_1 and S_2 are two positions of the X-ray source and π_1 and π_2 are images acquired at the X-ray detector respectively. The gantry angulation is changed by moving the source from S_1 to S_2 along R_S . For illustration, we consider a coordinate system with XY-plane as the object plane π' at some height above the table T . This object plane is an approximation of the position of the anatomy being imaged. The two obtained images are two acquisitions done in different geometry (table panning and rotation) and a plane, attached to the table. There exists an homography between these two images if we consider that the imaged objects lies entirely in this plane. This homography is an approximation for objects placed slightly above or below this plane. We consider that this is the case for the anatomy of interest.

The homography is very simply determined by taking 4 points in the considered planes thus computing their projection matrices. Thereafter, using the projection matrices we can derive the matrix H from these two projections of 4 points by simple linear computations. This computation can handle both table panning and gantry motion if the projection matrices are defined in a referential attached to the object plane. In this setting, for any point X in the object plane π' , the relation between the points X_1 and X_2 , in the images π_1 and π_2 respectively, can be given by,

$$\begin{aligned} X_1 &= H_{1\pi'} X \\ X_2 &= H_{2\pi'} X \\ X_2 &= H_{2\pi'} H_{1\pi'}^{-1} X_1 = H X_1 \end{aligned} \quad (4.4)$$

where H , $H_{1\pi'}$ and $H_{2\pi'}$ are the homographies. The computed homography H is used for transforming π_2 . As we consider the plane parallel to table top as object plane, the determined transformation is perfectly exact for points belonging to this plane and is inexact for points out of plane. The error grows as the distance to this plane increases. The error is smaller if the gantry angulation difference is smaller. For our task, we compute homography and transform all images in the reference sequence and fluoroscopic image stream to the reference image at first cardiac phase.

4.3 VOIDD algorithm

In this section, we first elaborate the general tracking framework of the proposed VOIDD algorithm (Section 4.3.1). We then explain (Section 4.3.2) how to extract the features (from the *reference sequence* and the *fluoroscopic image stream*), which are then used by the VOIDD algorithm.

Algorithm 1: VOIDD

Data: A fluoroscopic image stream \mathcal{F} and a reference sequence \mathcal{R}
Result: Track T_{VOI} representing detected vessel of intervention

```

1 Initialize  $\mathbb{T}$ , a dictionary of tracks ;
2 foreach image  $F$  in  $\mathcal{F}$  do
3    $(F, R) := \text{CreateImagePair}(F, \mathcal{R})$  ;
4    $\mathbb{P} := \text{ExtractFeaturePairs}(F, R)$  ;
   // feature pairs are ranked in decreasing order of matching
   // score
5   foreach  $\mathbb{P} \in \mathbb{P}$  do
6      $T_{\text{best}} := \text{nil}$  ;
7      $\Psi_{\text{TAC}}^{\text{best}} := \lambda_{\text{TAC}}$  ;
8     foreach  $T \in \mathbb{T}$  do
9       if  $(\text{TrackNotAssigned}(T, F))$  then
10         $\Psi_{\text{TAC}} := \text{TrackAssignmentCost}(T, \mathbb{P})$  ;
11        if  $\Psi_{\text{TAC}} < \Psi_{\text{TAC}}^{\text{best}}$  then
12           $T_{\text{best}} := T$  ;
13           $\Psi_{\text{TAC}}^{\text{best}} := \Psi_{\text{TAC}}$  ;
14        if  $T_{\text{best}} \neq \text{nil}$  then
15           $\text{AssignTrack}(T_{\text{best}}, \mathbb{P})$  ;
16        else
17           $T_{\text{new}} := \text{MakeTrack}(\mathbb{P})$  ;
18           $\text{AddTrack}(\mathbb{T}, T_{\text{new}})$  ;
19  $T_{\text{voi}} = \text{LongestTrack}(\mathbb{T})$  ;

```

4.3.1 General tracking framework

We aim to obtain the VOI by making a smart correspondence between the reference sequence \mathcal{R} and the fluoroscopic image stream \mathcal{F} . Therefore, we propose an algorithm, called VOIDD, that is able to simultaneously detect the guidewire tip in the fluoroscopic images and the VOI, that is the branch of the coronary vessels in which the guidewire is currently navigating. From a broader perspective, the algorithm consists of: i) detecting feature pairs from iso-phase image pairs; ii) grouping these feature pairs into tracks, a track being a sequence of feature pairs that are spatially consistent in time; iii) selecting the most relevant track as the detected vessel of intervention. A *feature pair* $\mathbb{P} = (C, P)$ is made of two corresponding curves and is extracted from an iso-phase image pair (F, R) . The first one, called a *tip candidate* (C), is extracted from the fluoroscopic image and it possibly corresponds to the guidewire tip in the fluoroscopic image F . The second one, called a *vessel path* (P) is obtained from the reference sequence \mathcal{R} and is a path in the coronary vessel tree that optimally fits the associated tip candidate C . The precise description of the VOIDD algorithm and of the feature pairs extraction is given in algorithm 1 and in Section 4.3.2 respectively.

VOIDD algorithm manages a dictionary \mathbb{T} of tracks, where each track T in \mathbb{T} is a sequence of feature pairs, with at most one feature pair per image of the fluoroscopic image stream \mathcal{F} . For each iso-phase image pair (F, R) , the algorithm optimally assigns each detected feature pair to one of the existing tracks. A cost called the *track assignment cost* (described in Section 4.3.3), is considered to optimally assign a feature pair to a track. If there exists no adequate track for a feature pair, a new track in \mathbb{T} is initialized. Once all the images in the fluoroscopic image stream are processed, the longest track (*i.e.* the track with maximum number of feature pairs) is selected as the VOI.

4.3.2 Feature pairs extraction

This section elaborates the extraction of the feature pairs, which are associations between images in any iso-phase image pair (F, R) . For each input image F in the fluoroscopic image stream an iso-phase image pair is formed. The function $CreateImagePair(F, \mathcal{R})$ of algorithm 1 returns such a pair (F, R) given an input image F and the reference sequence \mathcal{R} . First, we explain the extraction of centerline of the injected vessels to obtain vessel graph in Section 4.3.2.1. This is followed by the tip candidate extraction by morphological segmentation and thinning, explained in Section 4.3.2.2. Finally, we present in Section 4.3.2.3, the matching part to find the possible associations (the vessel paths) of the tip candidate in the vessel graph.

4.3.2.1 Vessel centerline extraction

This section describes how reference image R is transformed into sparse representation by representing extracted centerlines of the vessels in the form of a graph. The task of centerline extraction can be divided into two steps: i) vessel enhancement, and ii) vessel centerline segmentation. To enhance vessels, articles in the literature

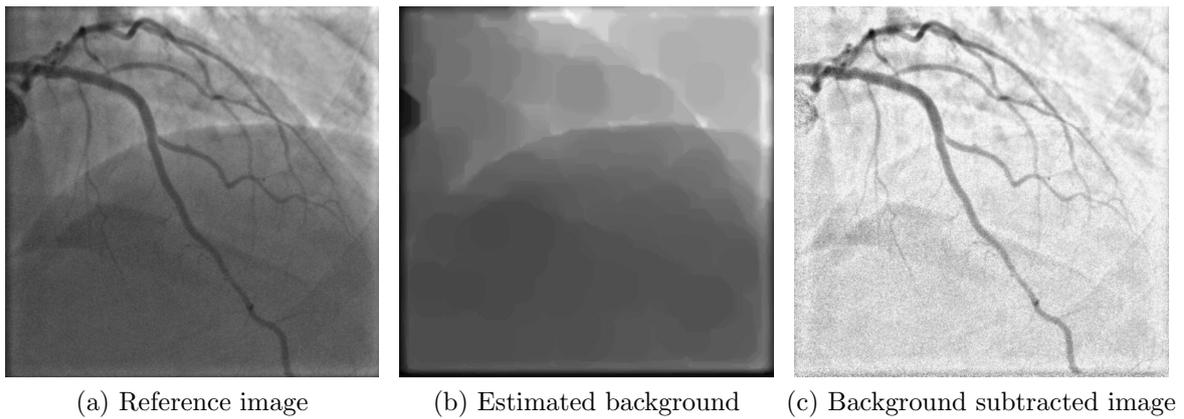


Figure 4.5: Background subtraction for contrasted vessels: (a) Native reference image, (b) estimated background with the help of morphological closing, (c) resulting background subtracted image.

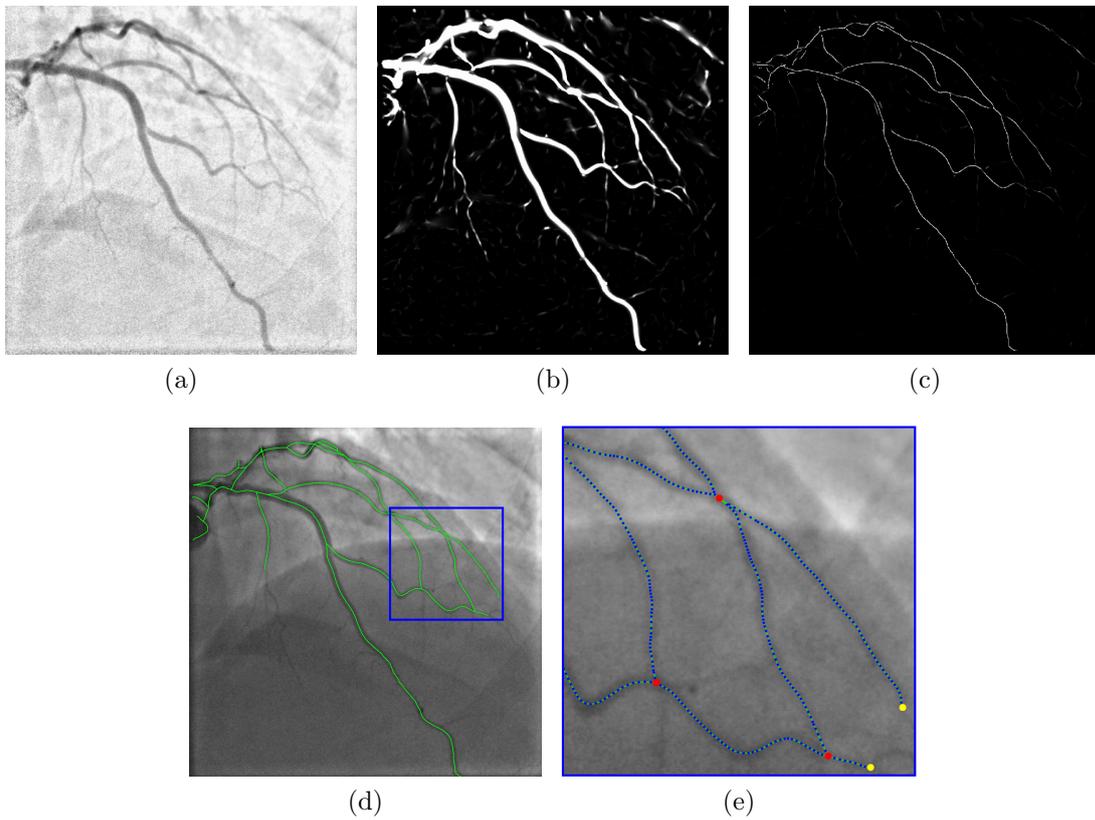


Figure 4.6: Vessel centerline extraction and graph representation: (a) Native reference image with injected vessels, (b) vesselness image, (c) result of Non-Maximum Suppression (NMS) applied to vesselness image, (d) extracted vessel centerlines superimposed on the native reference image and the region marked in blue is then zoomed in (e) where the vertices of different degrees of the vessel graph \mathcal{X} . Vertices of degree 1 in yellow, vertices of degree 2 in blue, vertices of degree 3 or more in red.

often estimate the anatomical background by morphological operations. Each image R in the reference sequence \mathcal{R} is preprocessed with a morphological closing to estimate the background image. This estimated background is used to remove the contrasted structures whose size is derived from the estimated maximum of diameter of coronaries in angiographic images, thus obtaining a background subtracted image as presented in Figure 4.5. However, some negatively contrasted anatomical structures like (ribs, spine and pulmonary structures) may persist in the background subtracted image. To enhance the vessels in the resulting background subtracted image an Hessian based technique [Krissian et al., 2000] is adapted which assumes that vessels are locally linear tubular structures. Since the coronary vessels have a full range of possible radii, the vesselness image obtained by [Krissian et al., 2000] contains vesselness measure at different scales as in the method of [Frangi et al., 1998]. Figure 4.6b shows the obtained vesselness image from the background subtracted image in Figure 4.6a.

In order to have a geometrical representation of the vessel structure, we propose a centerline extraction scheme to obtain vessel centerlines from the vesselness image. This centerline extraction includes a first step of Non-Maximum Suppression (NMS) that gives a skeletonized vesselmap with only pixels at the vessel center preserved. The resulting NMS image is segmented with a hysteresis thresholding which eliminates components formed of pixels with low vesselness values only. The pixels in resulting image (Figure 4.6c) from the hysteresis thresholding are organized in a graph structure \mathcal{X} , such that each pixel is a vertex of the graph [Benseghir, 2015]. According to this graph representation, the vertices of degree 1 correspond to the vessel extremities, the vertices of degree 2 corresponds to the vessel branches and higher degree vertices correspond to the bifurcation of the vessel tree (see illustration in Figure 4.6), the degree of a vertex being the number of edges incident to this vertex. We call a *vessel path*, any simple path in this graph \mathcal{X} , that is a sequence

$$P = \langle P_0, \dots, P_\ell \rangle, \quad (4.5)$$

such that (P_{i-1}, P_i) is an edge of \mathcal{X} for any $i \in \{1, \dots, \ell\}$ and,

$$P_i \neq P_j \quad \forall i, j \in \{1, \dots, \ell\} \text{ s.t. } i \neq j. \quad (4.6)$$

The integer ℓ is called the length of P and given any vessel path Q , we denote by $\#(Q)$, the length of Q . The set of all possible vessel paths in a graph \mathcal{X} is denoted by $\Pi_{\mathcal{X}}$. Hence, for each reference image R a graph \mathcal{X}_R is computed as described above, providing us with a representation of the vessel tree at each phase of the cardiac cycle.

4.3.2.2 Tip candidate extraction

A guidewire tip appears as a contrasted thin and elongated object (input image in Figure 4.3) in a fluoroscopic image. Several authors have worked on the task of segmenting the guidewire [Vandini et al., 2017; Chen et al., 2016; Wang et al., 2017, 2009]. For PCI application as in [Honnorat et al., 2012], the weak contrast of

the guidewire body makes the task very challenging. Therefore, we choose to rely on the more contrasted guidewire tips. In order to detect possible locations of the guidewire tips in a fluoroscopic image, we use the same image representation, the min tree, which was used for segmentation of empty catheter in Chapter 3. The detailed description of the min-tree has been made in the Section 2.2.6.1.

We pre-process the input image to remove the anatomical background with the help of morphological closing. As a brief review of min tree framework, for a grayscale image (pre-processed fluoroscopic image in this case) F defined on a domain \mathcal{D} and with values in range \mathbb{R} the *lower-level set of F with level h* in \mathbb{R} is

$$F_h = \{x \in \mathcal{D} \mid F(x) \leq h\} . \quad (4.7)$$

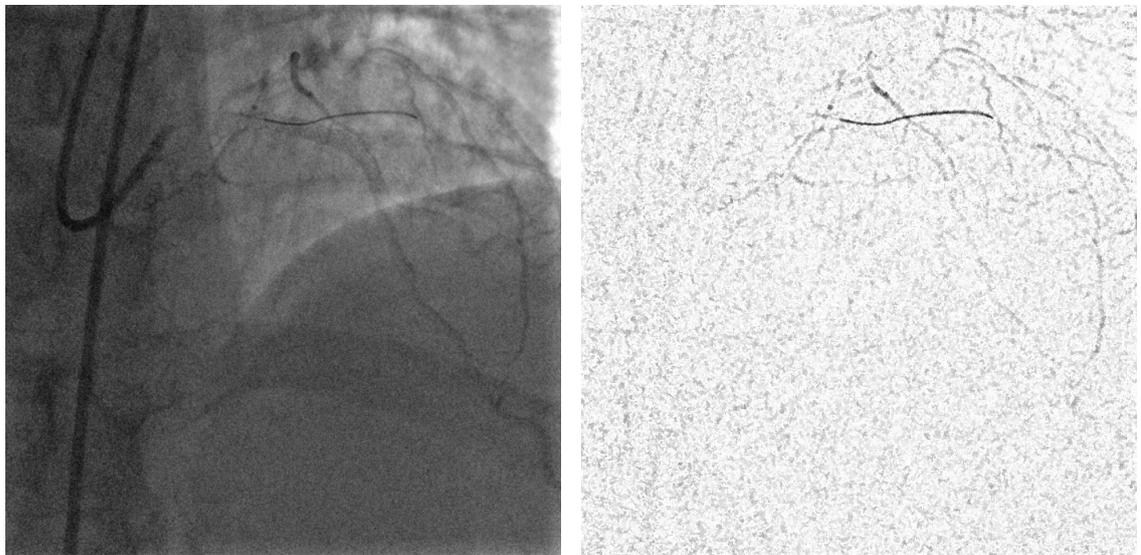
The image F when thresholded in an increasing order at every possible gray level in the range \mathbb{R} , yields a stack of nested (lower-) level sets. Each level set can be partitioned into connected components when the domain \mathcal{D} is structured as a pixel-adjacency graph (we consider 4-adjacency relation). Based on inclusion relationship, the structured set of all connected components is a directed tree called the min tree of the image F , denoted by \mathcal{T}_F . We assign to any connected component Z of the min tree \mathcal{T} , an elongation attribute $\mathcal{A}(Z)$ allowing us to characterize the guidewire tip as an elongated and thin connected component. More precisely, for any component Z , we set

$$\mathcal{A}(Z) = \frac{(\pi \times l_{max}(Z)^2)}{|Z|} , \quad (4.8)$$

where $|Z|$ is the cardinality of Z and $l_{max}(Z)$ is the length of the largest axis of the best fitting ellipse for Z . Since the guidewire tip is thin and long, the connected components corresponding to the tips have high value of attribute \mathcal{A} . A mere thresholding of the elongation \mathcal{A} is not sufficient, often giving other long and elongated (unwanted) objects like pacing lead and filled catheters. Indeed, these objects have higher elongation value than the guidewire tip. Hence, according to physical properties of the guidewire tip (recalled in Section 4.2.2), we set an upper bound value t_{max} on \mathcal{A} to the maximum possible elongation value of the guidewire tip, to ensure that extracted components contain guidewire tip. Even with this upperbound threshold keeping the most elongated components does not always lead to the desired tip.

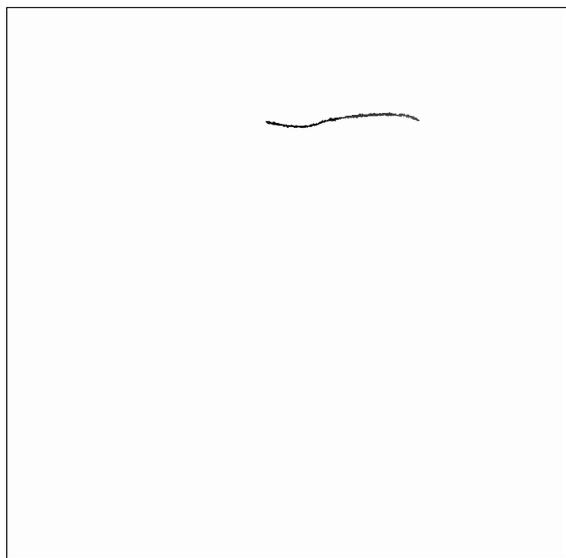
Therefore, we adopt the shaping framework defined in Section 2.2.7 that allows us to efficiently extract significant connected components. We apply shaping filters [Xu et al., 2016] with a height criterion: using the Min-tree $\mathcal{T}\mathcal{T}$ of the weighted graph $(\mathcal{T}, \mathcal{A})$ we preserve the nodes that are filtered in the first step and which have certain height in $\mathcal{T}\mathcal{T}$. The extracted components constitute the tip candidates.

We describe a segmented tip (wire-like object) by its centerline because we consider it as a 1-D object. This simplification to 1-D object aids in the matching step of feature pair extraction. Therefore, we perform skeletonization [Couprie and Bertrand, 2012; Saha et al., 2016] of the selected connected components to obtain the centerlines of the tip candidates. The medial axis (or skeleton) of a connected component is often sensitive to the noise on the boundary of the connected compo-

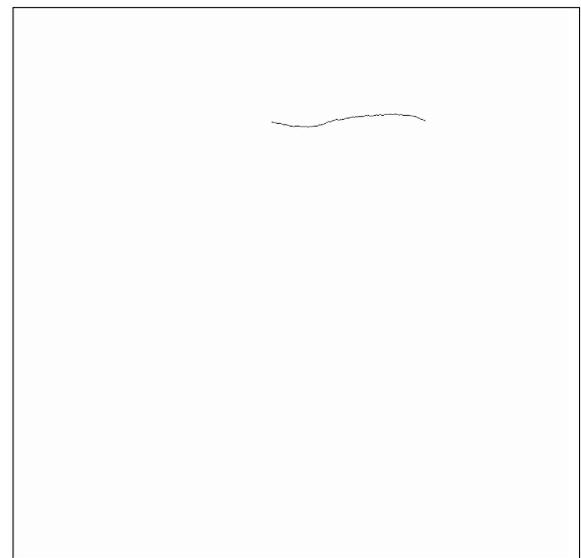


(a) Input image

(b) Background subtracted image



(c) Segmented guidewire tip



(d) Skeletonized guidewire tip

Figure 4.7: Tip candidate extraction: (a) Native fluoroscopic image, (b) background subtracted image, (c) result of guidewire tip segmentation, (d) segmented guidewire tip after skeletonization

ment, leading to a large number of (spurious) branches on the medial axis. Hence, the adopted skeletonization techniques include some regularization to prune the unwanted branches [Saha et al., 2016; Shaked and Bruckstein, 1998], allowing the resulting skeleton to contain only one branch and two ends. By its very definition, the medial axis of an object does not in general reach the border of the object, which results to a loss of geometrical information for our application. In order to recover this lost information, we extend the skeleton at the two ends along its tangents till the border of the connected component. Fig. 4.7 shows an example of the obtained centerline of segmented guidewire tip from the input image. This centerline of the tip candidate C is modeled as a series of points

$$C = \langle C_0, \dots, C_\ell \rangle \text{ s.t. } \forall i \in \{0, \dots, \ell\}, C_i \in \mathbb{R}^2 \quad (4.9)$$

where C_0 and C_ℓ denote the end-points of curve C . Such a series of points, when pairwise interpolated by straight lines, can be equivalently considered as a polygonal curve \mathcal{Q}_C defined by:

$$\mathbf{x} \in \mathcal{Q}_C \iff \begin{cases} \exists i \in \{0, \dots, \ell\} \\ \exists \alpha \in [0; 1] \end{cases} \text{ s.t. } \mathbf{x} = \alpha \cdot C_{i-1} + (1 - \alpha) \cdot C_i \quad (4.10)$$

4.3.2.3 Matching

An important step in the task of VOI detection is to designate possible associations of the guidewire tip inside the injected vessel. In the previous sections, we presented extraction of necessary information from an iso-phase image pair. We recall that for any iso-phase image pair (F, R) , the set \mathbb{C}_F is set of all tip candidates and the set $\Pi_{\mathcal{X}_R}$ is set of the vessel paths of the corresponding vessel graph \mathcal{X}_R . The matching step refers to building association from each tip candidate to one or more vessel paths that corresponds to the possible location of the tip candidate in the vessel graph. The task to be performed in this step is matching and aligning a tip candidate to the vessel graph to find corresponding vessel path. Therefore, the task reduces to a curve matching problem.

A curve-to-curve distance has to be defined to match and align two curves. [Alt and Godau, 1992] addressed the problem of measuring distance between two polygonal curves and proposed a continuous version of Fréchet distance. The method of Fréchet distance [Alt and Godau, 1992] based curve matching had been proposed to perform 3D 2D registration by [Benseghir et al., 2015]. In this case, we adapt the curve pairing algorithm of [Benseghir et al., 2015] to perform the task of matching as it takes into account the Fréchet distance and the shape similarity between two curves. It is required to define a curve-to-curve distance to compare any pair of elements from the two sets \mathbb{C}_F and $\Pi_{\mathcal{X}_R}$. We use a discrete version of Fréchet distance [Eiter and Mannila, 1994] as it takes into account the location and the order of points along the curves by establishing a monotone mapping from one curve to another. Imposed non-decreasing surjective mappings (reparameterization mapping) in computation of Fréchet distance takes into account the order of points along the curves. This order aids in curve pairing described below to give

scan direction along the curves. The shape similarity between two curves as defined in [Benseghir et al., 2015] is curve resemblance distance d_R , which is the residual Fréchet distance after 2D rigid registration between two curves. The resulting curve pairing distance $d_C(Q, Q')$ between any two polygonal curves Q and Q' is given by,

$$\begin{aligned} d_C(Q, Q') &= \beta \cdot d_F(Q, Q') + (1 - \beta) \cdot d_R(Q, Q') \\ \text{s.t. } d_R(Q, Q') &= d_F(T_{2D} \circ Q, Q') \end{aligned} \quad (4.11)$$

where $\beta \in [0; 1]$ controls the relative significance between the geometric and shape criterion. Given a tip candidate C , the above matching step requires the selection of every admissible vessel path P in the graph \mathcal{X}_R . In order to restrict computational complexity of search, we restrict the set of admissible vessel paths P to be in the neighborhood (search zone radius λ_E) of the tip candidate C extremities. Thus, we say that a pair $\mathbb{P} = (\langle C_0, \dots, C_m \rangle, \langle P_0, \dots, P_n \rangle)$ is a feature pair of the iso-phase image pair (F, R) whenever we have,

$$\begin{aligned} d(C_0, P_0) &\leq \lambda_E \quad \text{and} \\ d(C_m, P_n) &\leq \lambda_E \quad \text{and} \\ d_C(C, P) &\leq \lambda_C \quad , \end{aligned} \quad (4.12)$$

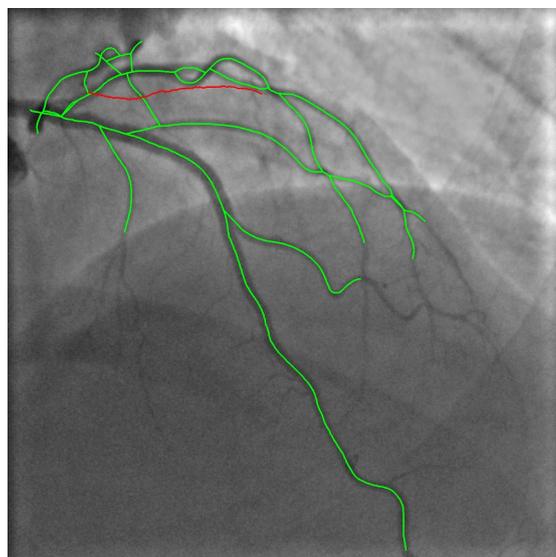
where d and d_C represent Euclidean and curve pairing distances respectively. After extracting all feature pairs of an iso-phase image pair (F, R) as defined by equation (4.12), we filter out redundant feature pairs. Two feature pairs $\mathbb{P}_a = (C_a, P_a)$ and $\mathbb{P}_b = (C_b, P_b)$ are redundant if,

$$C_a = C_b \quad \text{and} \quad d_H(P_a, P_b) \leq \lambda_H \quad (4.13)$$

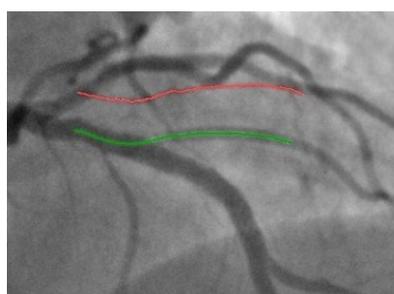
where d_H is the Hausdorff distance between two vessel paths. When two redundant feature pairs (C, P_a) and (C, P_b) are found, we keep (C, P_a) if $d_F(C, P_a) \leq d_F(C, P_b)$ or otherwise we keep (C, P_b) . This filtered set of feature pairs are sorted according to a curve resemblance distance (Equation 4.11) to prefer vessel paths with higher shape resemblance to the tip candidate. The filtered set of feature pairs is considered for each iso-phase image pair (F, R) and is denoted by $\mathbf{P}_{(F,R)}$. Thus, the filtered and sorted set $\mathbf{P}_{(F,R)}$ of feature pairs is what is returned by the function `Extract-FeaturePairs()` in algorithm 1. Figure 4.8 shows different feature pairs obtained for corresponding position of the guidewire tip shown on the iso-phase reference image.

4.3.3 Track assignment cost

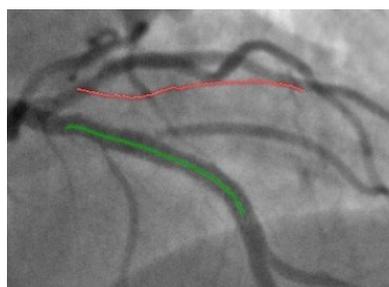
It is necessary to design a strategy to assign a new feature pair to the appropriate track. A track being a dictionary of feature pairs, the task of assigning a new feature to a track can be considered a measuring a kind of distance between the feature pair and the track. Intuitively, we may consider that for any new feature pair, the most appropriate track will minimize such distance.



(a)



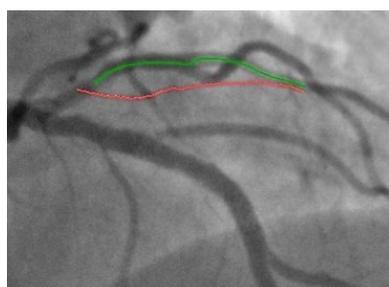
(b)



(c)



(d)



(e)

Figure 4.8: Feature pair extraction: (a) Tip candidate (red) in the iso-phase reference image with extracted centerlines (green), (b), (c), (d) and (e) corresponding feature pairs with tip candidate (in red) and vessel paths (in green)

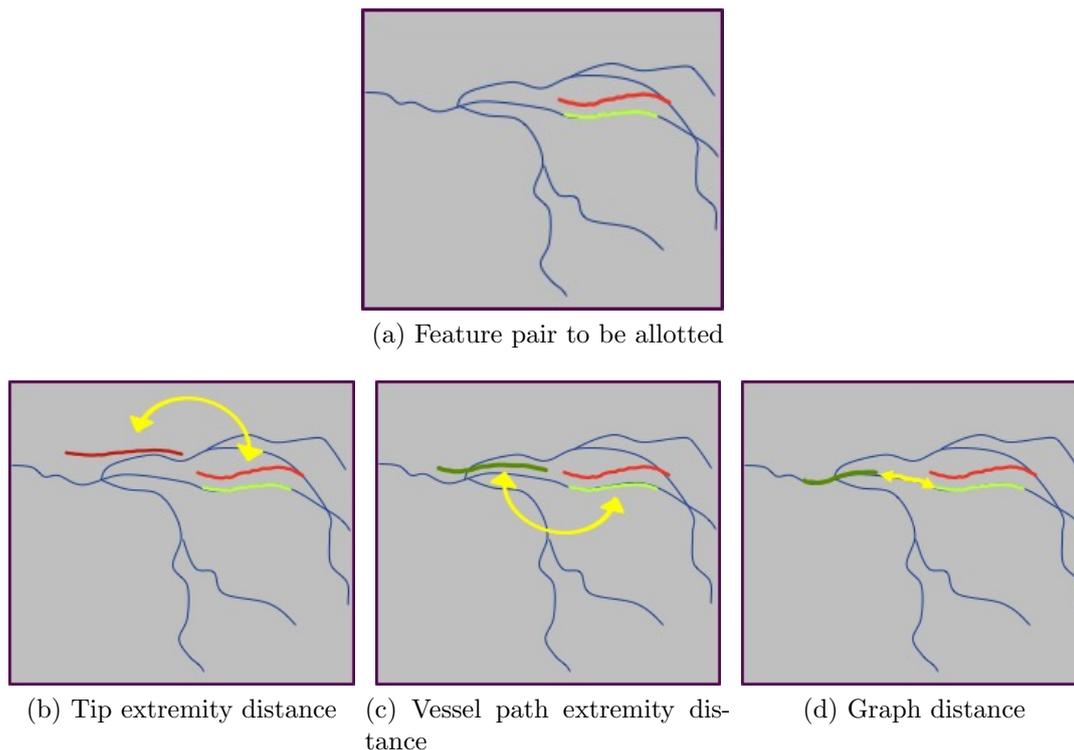


Figure 4.9: Illustration of track assignment cost: (a) shows the feature pair to be allotted, (b) is the tip extremity distance between tip candidate of the feature pair and the latest feature pair in considered track, (c) is the curve extremity distance between the corresponding vessels paths of the above two feature pairs, (d) is the graph distance between the vessel path of the feature pair and vessel path of latest iso-phase feature pair in the considered track.

The track assignment cost Ψ_{TAC} is computed as the cost to assign a proposed feature pair $\mathbb{P} = (C, P)$ to a track T . We design the track assignment cost with the spatial distance between the feature pairs and also the geodesic distance along the vessel graph. It is a combination of three distance values:

1. a tip extremity distance,
2. a vessel path extremity distance, and
3. a graph distance.

The two extremity distance values account for the geometrical shift between the two tip candidates and the two vessel paths respectively. The vessel path extremity distance and the graph distance preserve coherency in the tracks, whereas tip curve extremity distance avoids falsely detected guidewire tips to be added in the track.

The two extremity distance values are defined with the curve extremity distance which is the Euclidean distance between the extremities of two (non-oriented) curves. Given the positions of the guidewire tip (hence, feature pair) apparent due to navigation in the vessel, we consider the distance between two curves as mean

distance between their extremities. More precisely, given two curves C and C' , we define the *curve extremity distance between* $C = \langle C_0, \dots, C_m \rangle$ and $C' = \langle C'_0, \dots, C'_n \rangle$ by

$$d_E(C, C') = \frac{1}{2} \min_{(i,j) \in \{(0,n), (n,0)\}} \left\{ d(C_0, C'_i) + d(C_m, C'_j) \right\} \quad (4.14)$$

where $m = \#(C)$ and $n = \#(C')$. Then the third distance value is obtained with the graph distance $d_{\mathcal{X}}$ that is the mean length of the geodesic paths (in the graph \mathcal{X}) between the end-points of two given vessel paths in \mathcal{X} . Figure 4.9 illustrates with an example the three distances composing the track assignment cost. We transform these three distances with a non-linear normalizing function of the form $\nu(x) = 1 - e^{-\lambda x}$, so that they belong to the same range $[0, 1[$. The parameter λ of this normalizing function is set according to the length of the guidewire tip (recalled in Section 4.2.2).

Given a feature pair $\mathbb{P} = (C, P)$ and a (non-empty) track T , if T contains a feature pair at the same cardiac phase as \mathbb{P} , then the track assignment cost $\Psi_{\text{TAC}}(\mathbb{P}, T)$ is given by

$$\Psi_{\text{TAC}}(\mathbb{P}, T) = \frac{1}{3} \left(\nu(d_E(C, C')) + \nu(d_E(P, P')) + \nu(d_{\mathcal{X}}(P, P'')) \right), \quad (4.15)$$

where $\mathbb{P}' = (C', P')$ is the latest feature pair in the track T and $\mathbb{P}'' = (C'', P'')$ is the latest iso-phase feature pair in T . Otherwise, if the considered track T does not contain any iso-phase feature pair for a given feature pair \mathbb{P} , then the cost $\Psi_{\text{TAC}}(\mathbb{P}, T)$ is given by

$$\Psi_{\text{TAC}}(\mathbb{P}, T) = \frac{1}{2} \left(\nu(d_E(C, C')) + \nu(d_E(P, P')) \right). \quad (4.16)$$

The number of admissible vessel paths for a given position of tip candidate can be more than one. This may lead to track fragmentation, *i.e.* breaking of the longest track and creation of multiple tracks representing different parts of VOI. To limit this track fragmentation, we favor longer track by multiplying the track assignment cost by a track length factor l_T derived from the latest n processed images. We consider the number $|T|$ of feature pairs assigned to the track T . Then, the track length factor l_T is defined as $|T|/n$. In order to decide if a new track should be initialized (see line 17 of algorithm 1), we determine a track assignment cost threshold (λ_{TAC}) as the theoretically maximum possible value of track assignment cost to assign a feature pair to a track. The computation of λ_{TAC} is based on the length of the guidewire tip and the maximum observed guidewire speed.

4.3.4 Vessel of intervention refinement

The longest track contains the detected VOI for different cardiac phases. Figure 4.10(a,b) depicts all vessel paths of the longest track (T_{voi}) for one of the cardiac phases. All these vessel paths constitute a first proposal of the VOI for this cardiac phase as can be seen in Figure 4.10(a,b). This first proposed VOI often has

small branches, loops, hole, missing connections/cut vessel resulting from individual pairings of the tip candidates. We perform various morphological operations to post-process the union of these vessel paths in order to refine the automatically detected VOI at a given phase. We first apply a reconnection and hole-closing technique based on guided homotopic transform [Bertrand and Couprie, 2007], taking advantage of the prior knowledge on the VOI which is a simply connected object (*i.e.* made of a single connected component without any hole). More precisely, given the firstly proposed VOI \mathfrak{X} , we apply a guided homotopic thinning from a simply connected binary object consisting of the whole image with the following characteristics:

1. deletion of points in \mathfrak{X} are forbidden; and
2. points at higher Euclidean distance from \mathfrak{X} are deleted in priority.

The results of this step is illustrated in Figure 4.10(c,d). This is followed by skeletonization [Bertrand and Couprie, 2007] to reduce thick parts of the object resulting from the hole-closing. Thereafter, a branch pruning is performed [Postolski et al., 2013] to remove spurious branches of the skeleton and to obtain the longest section of the detected vessel of intervention. The resulting VOI after refinement, as shown in figure 4.10(e,f), does not contain any small branch, loop, thickness or disconnection.

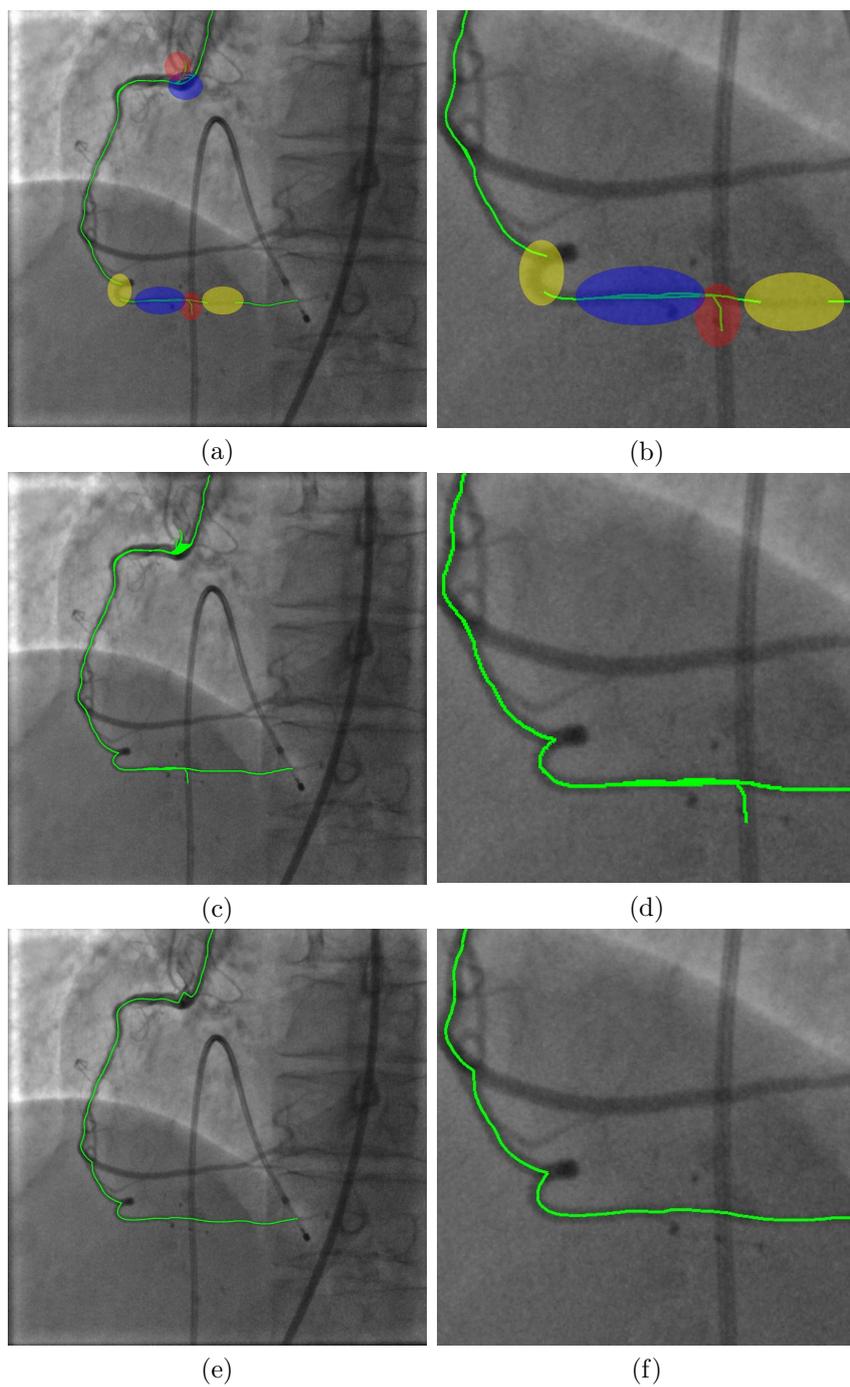


Figure 4.10: Vessel of intervention refinement, (a,b) VOI from longest track contains small branches (shaded red), loops (shaded blue) and hole/cut vessel (shaded yellow); (c,d) result after guided homotopic thinning step; (e,f) VOI after refinement (right). Images (b,d,e) on right shows zoomed sections of images (a,c,e) on left.

4.4 Dataset and evaluation measures

This section describes the dataset, the expert annotations (Section 4.4.1), and the three evaluation measures used to assess the results of the VOIDD algorithm. The tip detection evaluation measure (Section 4.4.2) analyzes the ability to detect the absence of tip in the beginning of the fluoroscopic image stream, the track evaluation measure (Section 4.4.3) analyzes the vessel paths in the feature pairs from the longest track, whereas the VOI evaluation measure (Section 4.4.4) estimates the ability to retrieve the VOI after refinement.

4.4.1 Dataset and expert annotations

The evaluation dataset consists of sequences collected in 15 different clinical cases including 9989 images from fluoroscopic image streams and 140 images from reference sequences (Innova 2100-IQ, GE Healthcare). These 15 clinical cases come from 14 patients. Each case contains 10 images in their reference sequence referring to the Φ different cardiac phases. The X-ray image size is 1024×1024 , and given the size of the pixels of the detector and the standard magnification, an image pixel corresponds to a structure of physical size about 0.14 mm. This dataset represents a variety of clinical cases, including different anatomies and vascular locations of the lesions, angulations of the gantry of the interventional system, the patient body mass indices (BMI), the image qualities (IQ) of the fluoroscopic images and the complexities of guidewire navigation. Table 4.1 illustrates the variation in this dataset of 15 clinical cases with respect to factors like the patient BMI, the VOI and the difference in gantry angulation between the reference sequence and fluoroscopic image stream. High patient BMI may exacerbate the image quality whereas large cranial angulation difference affects the appearance of vascular anatomy in two sequences. Some extreme cases of these factors may prevail in some patient data (patient L, M, N in Table 4.1). Based on the severity of extreme cases, we split the data for our evaluation study into two sets of 12 and 3 clinical cases.

The first set consists of 12 clinical cases named from A to K. These 12 clinical cases come from 11 patients, of which two clinical cases D_1 and D_2 refer to two different guidewires navigated in the patient D. For this patient, the first guidewire was navigated in distal left anterior descending artery whereas the second guidewire was navigated in a diagonal. Among the three extreme cases, which make up the second dataset, the patient M has high BMI (35.5 kg/m^2) leading to very low contrasted images in fluoroscopy and cine mode of interventional system (see Figure 4.16). On the other hand, the patients L and N have high gantry angulation difference between the reference sequences and the fluoroscopic image streams. Such difference affects the appearance of the (2D projected) vascular anatomy in the two image sequences. Therefore, the guidewire tip appears relatively far, from its actual position, in the vessel centerlines resulting in an incorrect matching. Generally, the physicians try to maintain the same gantry angulation for the acquisition of the two sequences. For these reasons, these three patient cases do not fulfill the conditions of the VOIDD workflow and fall outside the bound of hypothesis of the proposed

Clinical case	BMI kg/m ²	Name of VOI	Gantry angle difference(°)	
			LAO/RAO	CRA/CAU
A	29.20	Distal RCA	0.00	0.00
B	27.80	Distal LAD	0.80	6.40
C	19.60	RPD	2.50	0.60
D ₁	27.50	Distal LAD	0.00	0.00
D ₂	27.50	Diagonal	0.00	0.00
E	31.90	Diagonal	2.90	0.00
F	26.70	Distal LAD	2.70	1.90
G	23.50	Diagonal	4.00	0.20
H	32.00	Distal RCA	1.50	4.00
I	30.60	Distal LAD	3.10	0.80
J	31.50	LCX	11.70	4.30
K	34.80	Distal RCA	25.00	0.20
L	(NA)	Distal RCA	11.30	21.80
M	35.50	LCX	0.00	2.60
N	21.70	Proximal LAD	0.20	14.10

Table 4.1: Variation across the dataset of 15 patient sequences with three extreme cases. The different names of VOI are: RCA, right coronary artery; LAD, left anterior descending; RPD right posterior descending artery; LCX, left circumflex artery

method. A discussion on these three patient cases is presented in the next section to assess the proposed method on limit cases.

Each fluoroscopic image stream contains two different types of images: 1) images where guidewire tip is absent in the field of view at the beginning of the stream, 2) images where guidewire tip is visible in field of view and is navigating in the vessel of intervention. There may be images (shortly before the end of fluoroscopic image stream) where the guidewire tip is fixed at the distal end of the VOI. An expert user annotated the centerline of the branch of the artery navigated by the guidewire tip in the corresponding reference sequences of these patients. The expert annotations were marked by a single expert user using a semi-automatic software by selecting points of the vessel centerlines in the reference image extracted by the method presented in Section 4.3.2. However, it is not possible to define the exact corresponding location of the distal end of the guidewire tip in the vessel centerline because there is no visual information to determine the motion between the fluoroscopic and cine images. Therefore, a visually approximate location in vessel centerline image is decided as distal end, ensuring that this distal position is sufficiently away from the lesion. Thus, the expert annotations consist of the centerlines of 150 VOIs from 15 clinical cases. Each expert annotated centerline is modeled as a discrete polygonal curve, referred to as *GT* in the following sections. These centerlines are marked for each image of the reference sequence, representing different cardiac phases.

4.4.2 Tip detection evaluation measure

In order to evaluate the efficacy of the algorithm to identify the arrival and navigation of a guidewire (tip) in a fluoroscopic image stream, we analyze the robustness of VOIDD algorithm to detect VOI in images with no guidewire tip in field of view (FOV) in the beginning of the fluoroscopic image stream. In such sequence (of m images), we consider the length n of the longest track before the arrival of guidewire tip in FOV. The *tip detection accuracy* is computed as, $100 * (m - n) / m$. This evaluation measure conveys the ability of the algorithm to detect the arrival of the guidewire which in turn will help for automatic documentation of the PCI procedures.

4.4.3 Track evaluation measure

The longest track from the VOIDD algorithm includes multiple feature pairs, each of them indicating the location of the detected tip in the fluoroscopic image and the matched vessel path in the corresponding iso-phase reference image. Here, we assess the correctness of each individual feature pair (C, P) by considering the distance between the vessel path P of the feature pair and the expert annotation GT . We consider the following *target registration error (TRE)* between P and GT given by,

$$TRE(P) = \frac{1}{n} \sum_{i=1}^{i=n} \min_{\forall j \in \{1, \dots, m-1\}} \left| d\left(P_i, GT(j, j+1)\right) \right| \quad \left| \quad m = \#(GT), n = \#(P) \right. , \quad (4.17)$$

where $GT(j, j+1)$ refers to the segment between point GT_j and GT_{j+1} and d refers to point to segment distance converted to mm using typical known patient level pixel size (0.14 mm/pixel in our case). If the tip is correctly paired to the VOI then this TRE is governed by the usual small difference between the estimated centerline and the expert marked vessel centerline. The algorithm chosen feature pair is considered as a *correct detection* if the corresponding TRE is less than 0.5 mm. If the TRE is more than 0.5 mm, we consider that we have a *wrong detection*. If the input image contains a guidewire tip, but the algorithm does not provide any detection, then the TRE cannot be computed and a *missed detection* is reported. In Section 4.5, the results report the percentage of correctly detected feature pairs with respect to the total number of images with guidewire tip in the fluoroscopic image stream. Similarly, the wrong and missed detection are also reported as percentages of total images.

4.4.4 VOI evaluation measure

The track evaluation measure evaluates the feature pairs in the longest track whereas the VOI evaluation measure, presented in this section, analyzes the VOI after the refinement step (selected VOI). The selected VOI is modeled as a discrete polygonal curve. This evaluation method computes *VOI retrieval rate*, which is the fraction of expert annotated vessel of intervention correctly retrieved in the automatically selected VOI. The *VOI retrieval rate* is the ratio of the length of the correctly

Table 4.2: Performance of VOIDD algorithm on 12 clinical cases of 11 patients

Clinical case	Total images	Images without tip	Tip detection rate	Track evaluation measure		
				Correct detection	Wrong detection	Missed detection
A	320	156	98.72%	96.34%	0 %	3.66%
B	878	172	99.42%	89.24%	5.10%	5.67%
C	713	264	98.48%	92.87%	2.45%	4.68%
D ₁	302	98	97.96%	90.20%	2.45%	7.35%
D ₂	354	0	(NA)	61.58%	5.65%	32.77%
E	691	257	99.61%	83.87%	1.38%	14.75%
F	382	218	94.04%	95.17%	0.61%	1.22%
G	555	209	99.04%	83.53%	9.83%	6.65%
H	1002	546	99.45%	55.26%	1.32%	43.42%
I	781	366	100.00%	96.63%	2.65%	0.72%
J	710	248	99.60%	80.09%	6.28%	13.64%
K	1161	543	100.00%	92.72%	4.53%	2.43%
			99.05%	84.09%	4.21%	11.65%

detected part of the VOI over the length of the expert annotated VOI. The correctly detected part of the VOI contains the segments of the selected VOI which have their two extremities at a distance less than 0.5 mm from the expert annotated VOI. Each patient sequence has a selected VOI corresponding to each cardiac phase (number of cardiac phases being $\Phi = 10$). Therefore, the VOI retrieval rate for a patient is represented as the median, minimum and maximum of these Φ values.

4.5 Results

This section reports the performance of the VOIDD algorithm to detect vessel of intervention and assesses its potential to identify the arrival of a guidewire tip in the fluoroscopic image stream. Table 4.2 depicts the performance of VOIDD algorithm according to the tip detection evaluation measure and the track evaluation measure. The images in fluoroscopic image streams without guidewire tip are referred as images without tip in these tables. According to the tip detection evaluation measure, VOIDD algorithm robustly detects the tip with weighted mean detection rate of $99.05 \pm 1.61\%$ throughout the 12 clinical cases. Looking at this tip detection accuracy of VOIDD, we can use it to automatically detect the arrival of the guidewire tip and eventually for automatic documentation of PCI. The tip detection rate for the image stream D₂ (navigation of second guidewire) cannot be computed as these images already contain one guidewire tip navigated during the D₁ image stream.

According to the track evaluation measure in Table 4.2, the vessel paths in the longest tracks are correctly detected with an average accuracy of 84.09%. This

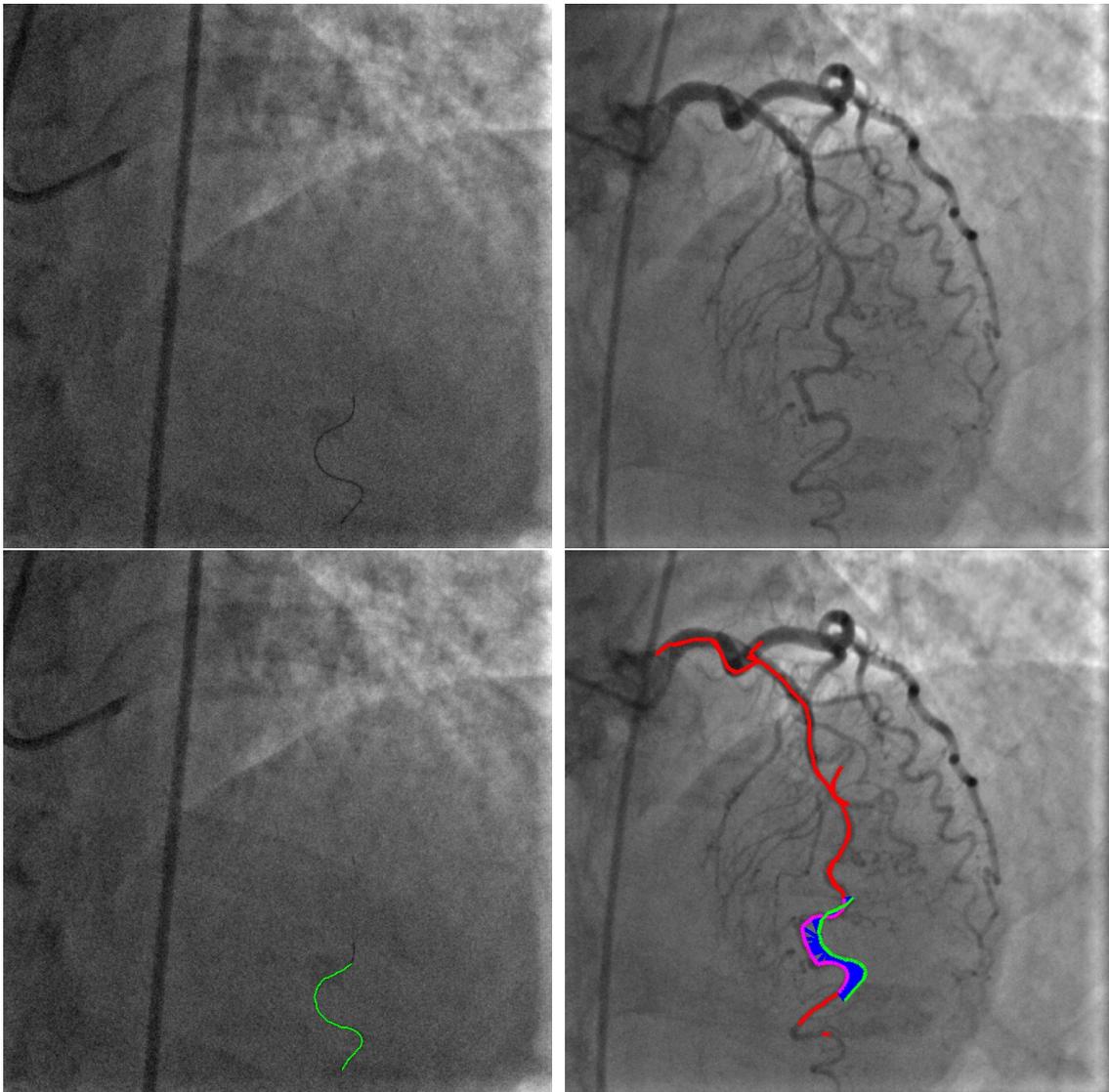


Figure 4.11: Correct detection example: (top) iso-phase image pair *i.e.* fluoroscopic image (left) with guidewire tip and the reference image at same cardiac phase; (bottom) segmented guidewire tip in fluoroscopic image and (on right) the result of matching (feature pair) and its location in the longest track.

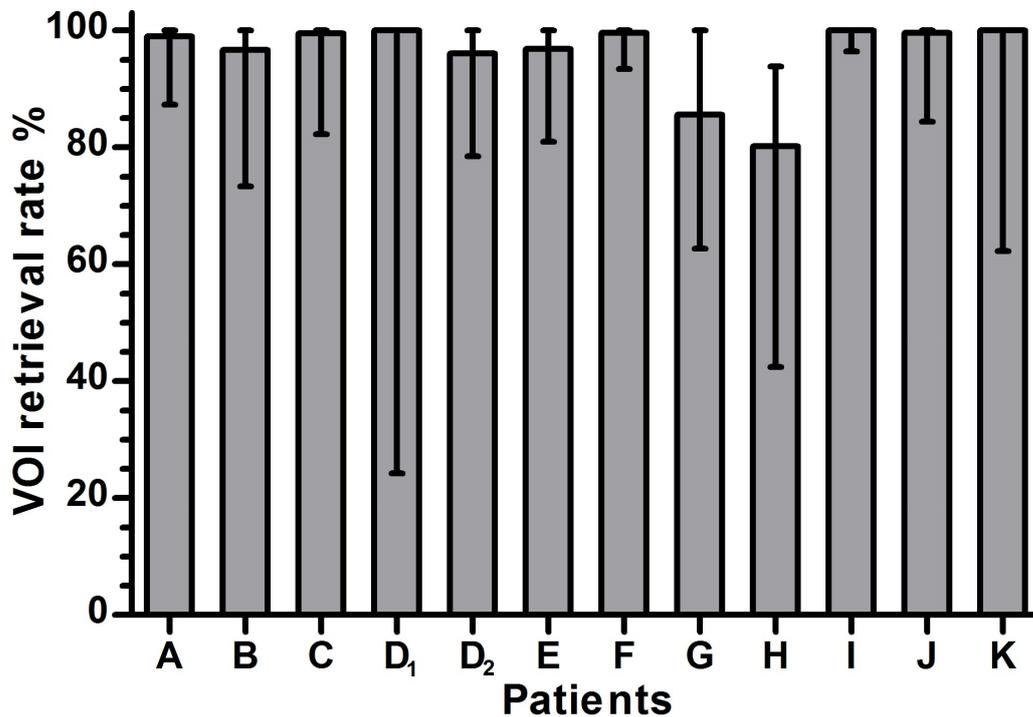


Figure 4.12: VOI retrieval rate

value is the mean correct detection rate weighted with the number of images with guidewire tip for corresponding image streams. Figure 4.11 illustrates an example of correctly detected vessel path in longest track for a given iso-phase image pair. Similarly, in these 12 clinical cases, the (weighted) mean wrong detection and the mean missed detection rates are 4.21% and 11.65% respectively. A non-zero wrong detection rate means that some vessel paths in the longest track do not correspond to the vessel of intervention. These wrong detections may refer to matching tip candidate to bifurcating branches of the vessel of intervention or due to the fact that sometimes the guidewire is pulled back by the operator after it entered by error in a non-targeted branch. The reporting of a missed detection implies non-assignment of guidewire tip to the longest track and the reasons for this being segmentation failure of guidewire tip or too large gantry angulation difference between reference sequence and the fluoroscopic stream. Failure of guidewire tip segmentation may occur due to: 1) blurred appearance of the tip because of its sudden movement imposed by cardiac motion, 2) reduced visibility of the tip caused by small contrast media injection to guide the navigation. Gantry angulation difference between the reference sequence and fluoroscopic image stream renders the guidewire tip away from the projected vessel centerlines, thus leading to missed detection. As compared to wrong detection, missed detection have a lesser impact on the ability to retrieve the vessel of intervention after refinement.

This 12 clinical cases (11 patients) dataset is further evaluated according to the

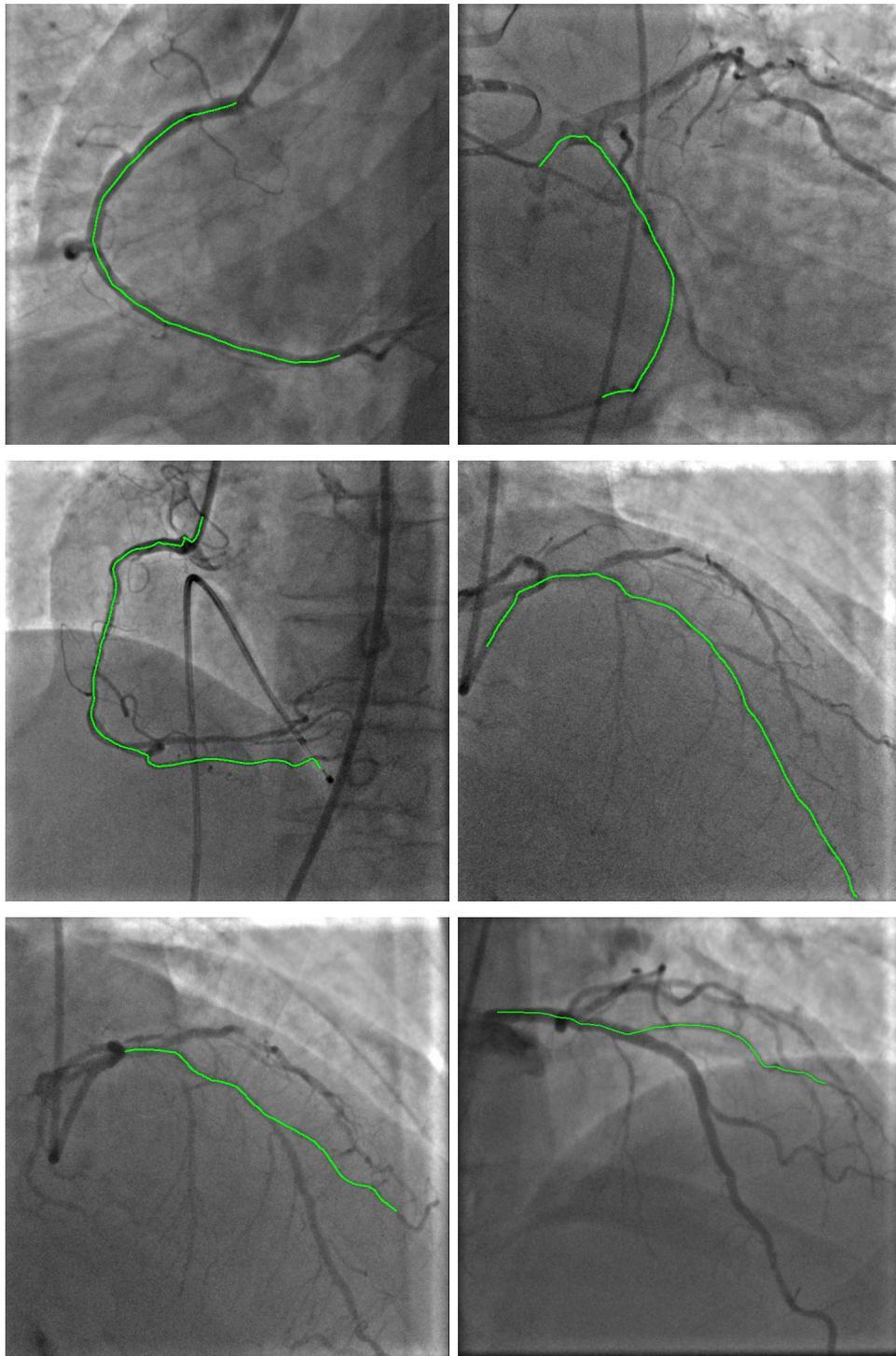


Figure 4.13: Detected VOI in 6 patient sequences

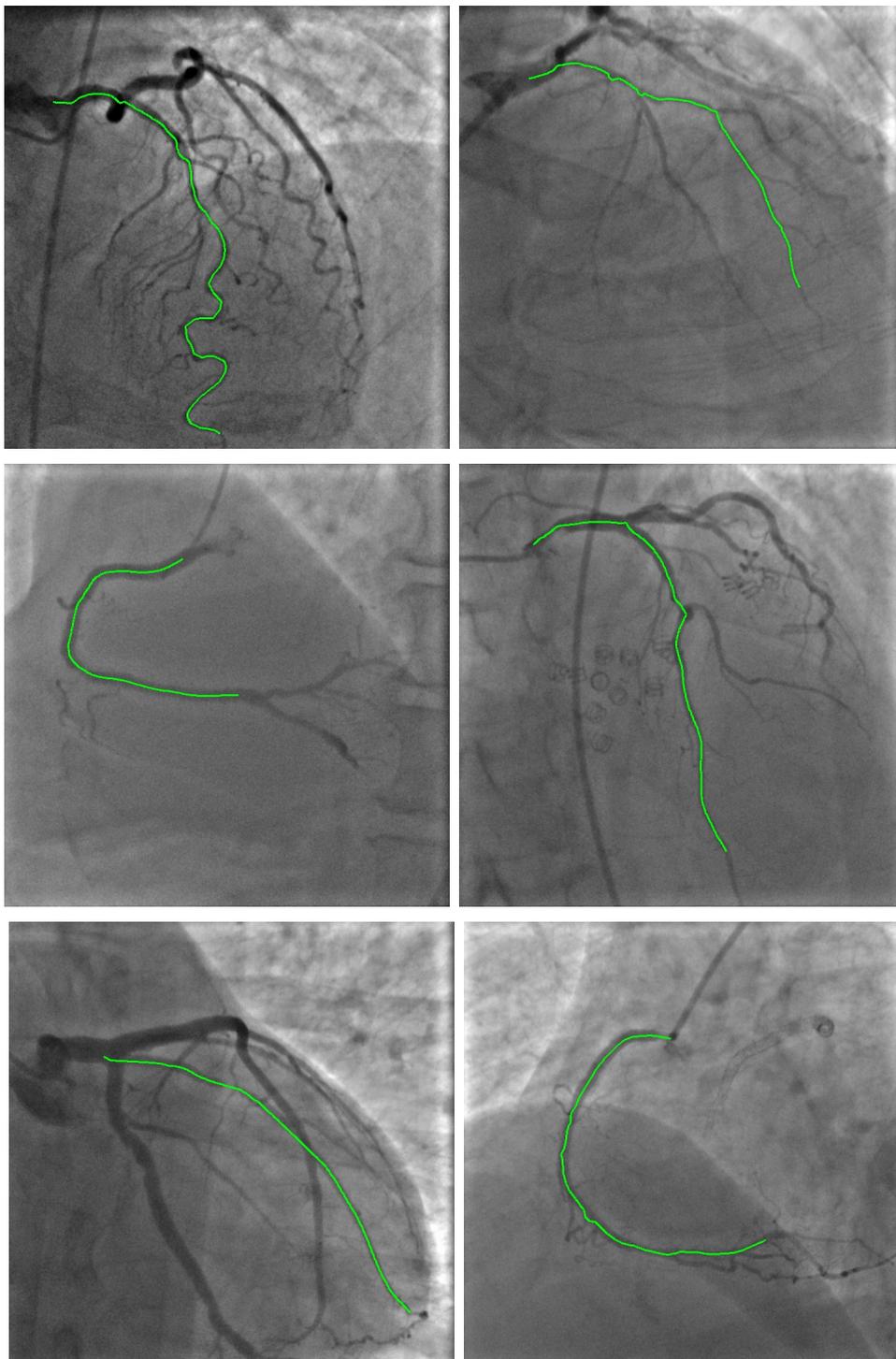


Figure 4.14: Detected VOI in 6 patient sequences

Table 4.3: Performance of VOIDD algorithm on second dataset of 3 clinical cases

Clinical case	Total images	Images without tip	Tip absence detection rate	Track evaluation measure			VOI retrieval rate
				Correct detection	Wrong detection	Missed detection	
L	551	243	99.18%	62.01%	11.04%	26.95%	74.08%
M	1060	202	99.50%	47.32%	11.31%	41.38%	45.67%
N	529	350	99.56%	20.67%	2.65%	76.54%	26.64%

VOI evaluation measure as shown in Figure 4.12, where the bars refer to median, minimum and maximum values of the VOI retrieval rate. For each case, these statistics are computed on the VOI retrieval rates for different cardiac phases ($\Phi=10$). The mean VOI retrieval rate for these 12 cases is 93.22%. This rate is significantly higher than the correct detection rate in track evaluation measure (average of 84.09%). This means that even if the algorithm does not completely succeed to correctly detect the tip in some images, the vessel of intervention may be correctly retrieved overall. The detected VOI for these 12 patient sequences at one of the cardiac phases is shown in Figures 4.13 and 4.14. The detected VOI for all (10) phases of patient E is shown in Figure 4.15, where Figure 4.15f depicts the phase where the VOI retrieval rate is 80.97% (lowest for this patient). The refinement step described in Section 4.3.4 is clearly helping here. In spite of the higher missed detections rates of sequences D₂ and H, the vessel of interventions were well retrieved after refinement, the mean VOI retrieval rates being 92.7% and 79.26% respectively. On further analysis it was observed that the detected VOI after refinement contained the lesion in all the cardiac phases of the 12 cases. Thus, we can claim 100% lesion detection for the algorithm. Intuitively, the inability to detect the VOI near the distal end of the vessel, leads to smaller VOI retrieval rate.

As reported in in Table 4.3, high BMI of the patient M leads to very low image quality (low-contrasted images). Figure 4.16 demonstrates difference in image quality for images with injected vessels, where the low image quality image belongs to patient M. Such low IQ may hinder the guidewire tip segmentation and vessel centerline extraction, eventually leading to a failure to detect the vessel of intervention. On the other hand, high gantry angulation differences between the reference sequences and the fluoroscopic image streams in the patients L and N have led to poor performance of the method. In these cases, the guidewire tip appears relatively far from vessel centerlines resulting in incorrect matching. Despite of these conditions, the mean tip detection rate for these three patients is 99.42%. We considered interesting to assess the performance of these cases because these two types of challenging conditions, high BMI and difference of angulation may be detected automatically and as such an application relying on our algorithm may automatically reject the case avoiding to deliver inconsistent results.

The parameters involved in various stages of the algorithm *e.g.* tip candidate extraction or track assignment cost were designed based on the physical properties of guidewire tip, permissible speed of advancement of guidewire. Current implementation runs in average 0.24 seconds/image for tracking and 1.12 seconds/patient for

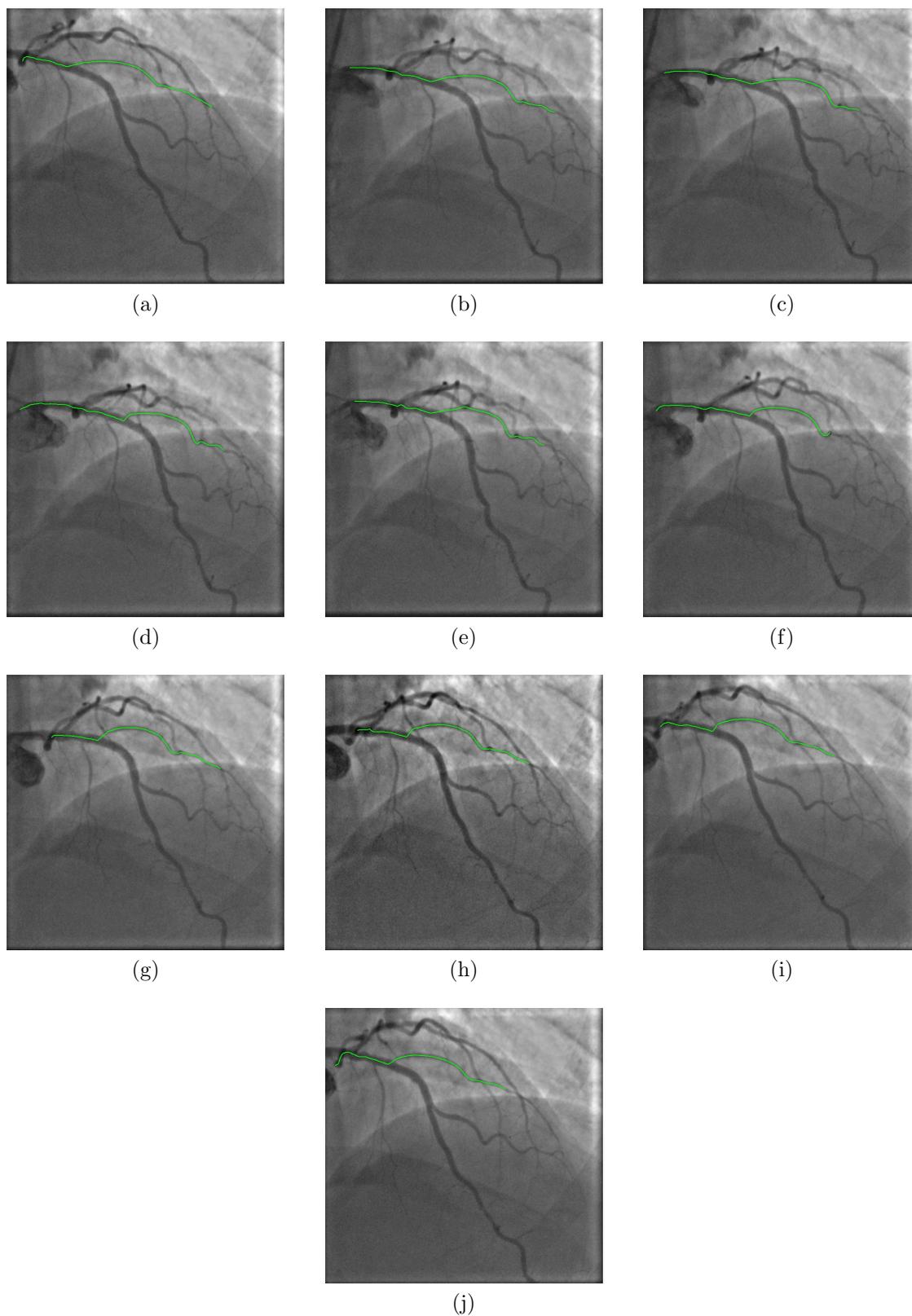


Figure 4.15: Detected vessel of intervention in 10 cardiac phases of patient E.

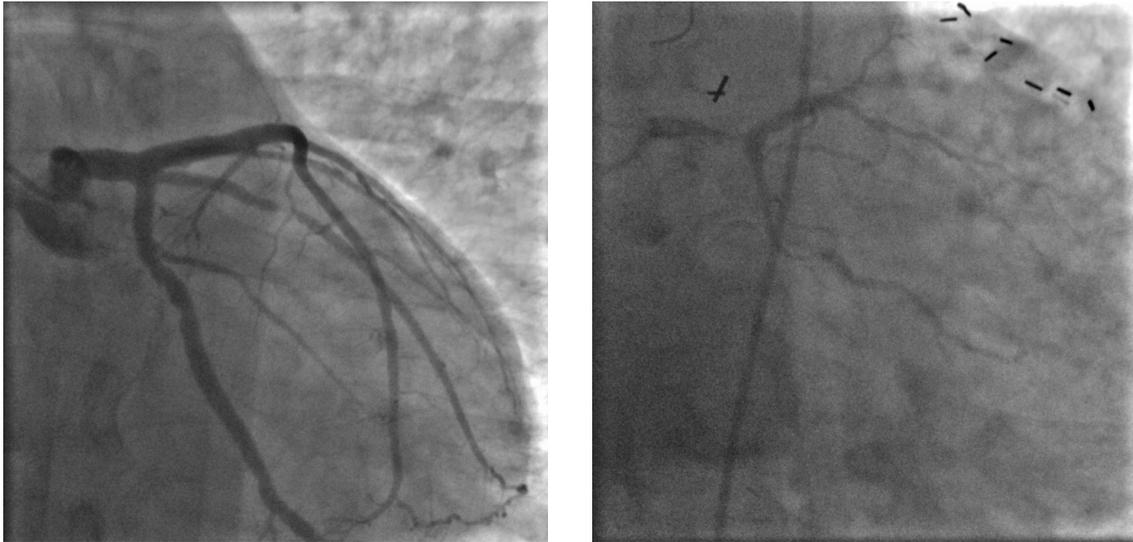


Figure 4.16: Image quality (IQ) demonstration: normal IQ reference image (on left) and low IQ reference image (on right).

the VOI refinement on a Intel Core i7 cadenced at 2.80 GHz. A video is available as supplementary material¹.

4.6 Conclusion and future work

We proposed in this chapter a framework to determine the vessel of intervention in fluoroscopic images during the PCI procedures. We also demonstrated the segmentation of the guidewire tip and the accuracy of detection of navigation of tip. We evaluated the accuracy and robustness of the algorithm for these two abilities. Encouraging results have been obtained with mean VOI retrieval rate of 93.22% and mean tip detection accuracy of 99.05%, proving robustness over different patient and imaging conditions. We believe this tracking algorithm can be extended to other applications like tracking IVUS probes. Also, unerringly efficient tip detection accuracy in all the 15 sequences, including the low IQ cases, proves the ability of the method to determine the arrival of the guidewire in the field of view. This algorithm has the potential to be part of the software embarked by X-ray imaging systems and capable of automatically monitoring the successive steps of the procedure in view of continuously adapting the system behavior to the user needs. This framework opens gates for development of varied applications like automatic documentation of the procedures, predicting the beginning of the navigation phase, improving the stent enhancement algorithms among others. Future work includes investigating the detection of the other major tools (marker balls, balloons) taking aid of detected vessel of intervention and their integration into a semantic model of the procedure.

¹<https://voidd-demo.github.io>

Conclusion and perspectives

Contents

5.1 Conclusion	111
5.2 Perspectives	113
5.2.1 Vessel of intervention for semantic applications	114
5.2.2 Cathlab of the future and Robotics in PCI	115

5.1 Conclusion

This PhD thesis in the field of image processing falls into the context of curing coronary heart diseases with minimally invasive procedures. Percutaneous coronary intervention is one such procedure which is performed under the control of real-time X-ray images in an interventional suite. During the procedure, a significant amount of contrast agent is used to visualize the vasculature of the patient under the X-ray images. The ionizing radiations are harmful to the patient and the clinician, whereas the high amount of contrast agent used during the procedure may affect the renal functions of the patient. Therefore, it is important to optimize the amount of contrast agent and X-ray radiations during the PCI procedure. We proposed in this thesis the concept of *PCI procedure modeling*, which intends at optimizing the behavior of the interventional machine according to phases of the procedure. As a preliminary step in this direction we elaborated the concept and described an approach based on the segmentation of interventional tools to detect the phases of the procedure. The concept of PCI procedure modeling was developed with the clinical objective of optimizing the use of the interventional machine to reduce the amount of X-ray dose and contrast under the constraint of not affecting the workflow of the procedure.

However, the tasks of segmentation and extraction of semantic information from X-ray images present some significant challenges as discussed at different parts in the thesis. We briefly summarize the different challenges here:

- Low exposure modes of X-ray system;
- Limited contrast of interventional tools;
- Cardiac and breathing motions;
- Multiple overlapping tools apparent in 2D projective X-ray images.

To this extent, we designed, studied and validated a segmentation technique to detect empty catheter, a low contrasted tubular structure, in fluoroscopic images. Furthermore, we proposed an automatic vessel of intervention detection method thus creating a semantic link between the two different phases of the procedure. These developments have been published and presented in conferences and one application of PCI procedure modeling has been patented (refer publication list [A](#)).

As it has been discussed in Chapter 2, though some articles in literature demonstrated segmentation of some interventional tools in X-ray images, they were intended for particular applications and could not be extended for segmentation of other tools. With the aim of developing generic framework, we propose a unique tool for segmenting different interventional tools. The tree-based shape spaces used in this thesis are based on component trees from mathematical morphology. It is possible to involve arbitrarily large and heterogeneous sets of knowledge in segmentation processes by associating to each node of the component-tree several attributes. This can lead to very accurate descriptions of the different structures to be segmented. In our case, these different structures are number of interventional tools. Throughout the thesis, we make an extensive use of the tree of lower-level sets of image, called the min tree, to segment interventional tools which appear as dark structures in the X-ray images.

We presented the segmentation of empty catheter in fluoroscopic images in Chapter 3. The catheter, being a commonly visible landmark and first tool to appear in the field of view, its segmentation is an important step towards PCI procedure modeling. The empty catheter (without contrast agent) appears as a low contrasted structure with two parallel and partially disconnected edges in fluoroscopic images. In other words, the dark structures (edges) belonging to the empty catheters are not connected in the image space. Therefore, we filter the min tree to extract curve blobs, which are tentative regions of the edges of the catheter. We then propose a novel structural scale-space, a hierarchy built on the blobs. The deep connected components in this hierarchy are explored to detect empty catheter. We design appropriate validation methodologies to quantify and evaluate the performance of the segmentation algorithm. The experimental results are presented on a database of 1279 fluoroscopic images and 630 record images. For fluoroscopic images, we assess two versions of our automatic algorithm with and without temporal feedback, whereas, for record images we analyze the performance with temporal feedback. We demonstrate the following results:

- In case of fluoroscopic images, the mean precision and recall for our method improves from $62.40 \pm 18.86\%$ and $55.84 \pm 18.2\%$ respectively, without temporal feedback, to $83.85 \pm 23.29\%$ and 67.87 ± 19.53 respectively, with temporal feedback.
- In case of record images, the mean precision and recall values are $88.2 \pm 13.54\%$ and $71.5 \pm 12.78\%$.

This method for empty catheter segmentation can also be extended for the segmentation of other low contrasted structures.

The fully automatic tracking algorithm, VOIDD, depicted in Chapter 4, is designed to address the problem of vessel of intervention detection. During the PCI procedures, the vessel of intervention is the significant region of interest in the X-ray field of view where different interventional tools are navigated in and out. The proposed framework combines information from a vessel image with contrast agent injection and fluoroscopic images acquired during guidewire tip navigation where the VOI is not visible. As a part of this study, we demonstrated a robust guidewire tip detection method. Therefore, the VOIDD algorithm has been demonstrated for two abilities:

- detection of vessel of intervention, and
- detection of guidewire navigation phase.

We propose validation methodologies and we evaluate the accuracy and robustness of the VOIDD algorithm for these two abilities. For the task of VOI detection, the mean VOI retrieval rate obtained is 93.22%. Encouraging results are obtained with mean guidewire navigation phase detection rate (also called tip detection rate) of 99.05%. The evaluation is performed on 15 cases from 14 patients with varying imaging conditions, thus allowing to evaluate robustness of the algorithm. One of the most interesting observation is that even though the VOI retrieval rate is 93.22%, the detected VOI contained the coronary lesion in 100% of the cases. Such promising results prove the potential of the algorithm to aid the detection of other interventional tools like balloon markerballs and balloon which will arrive at the coronary lesion at the later part of the procedure. We discuss these in the next section on perspectives.

5.2 Perspectives

In interventional cardiology, the concept of PCI procedure modeling opens vast perspectives in the direction of cathlab of the future. Though it may seem to be an ambitious claim, we will try to list perspectives step by step to strengthen our claim. To this extent, we start with some preliminary works that has been explored without being significantly developed to be part of this PhD thesis. However, it seemed important to us to provide evidence of the potential next steps making use of the detected vessel of intervention (VOI).

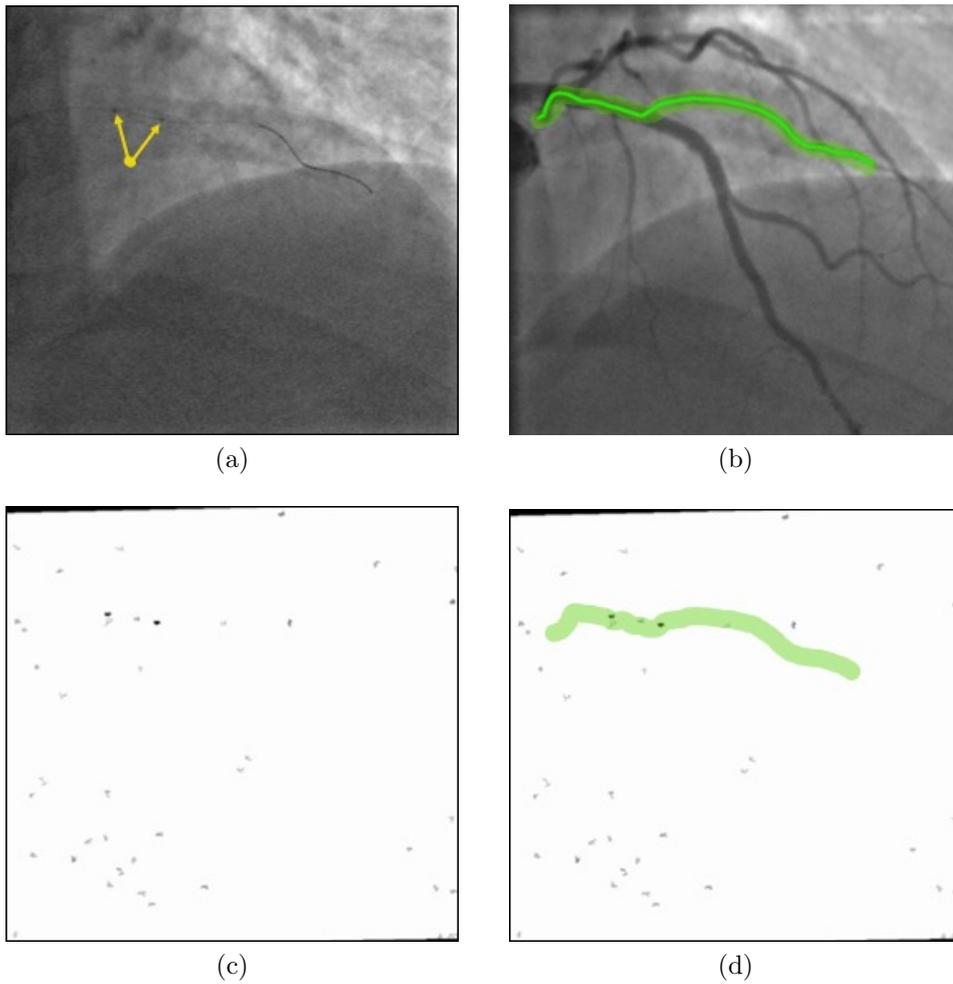


Figure 5.1: Markerball segmentation with semantic information based on vessel of intervention

5.2.1 Vessel of intervention for semantic applications

The VOIDD algorithm proposed in Chapter 4 returns a region of interest, *i.e.* a vessel of intervention, at the end of the guidewire navigation phase of the PCI procedure. This is followed by the markerballs positioning phase where a (stent mounted) balloon with markerballs is navigated to land at the location of lesion. As the name suggests, markerballs appear as pair of small dark contrasted circular structures in fluoroscopic images. An example in Figure 5.1a shows a frame of markerball navigation sequence. The corresponding vessel of intervention detected is shown in Figure 5.1b. As previously mentioned, we use the same min tree structure and filter the nodes with circularity and area attributes to select candidate nodes for markerballs. The filtered image as in Figure 5.1c contains many other connected components which are candidates for markerballs. This is mainly because of the low SNR of fluoroscopic images. However, we may use the detected VOI as region of interest to select and track marker balls among these various candidates.

In case of above example, the corresponding image by fusing the filtered markerball candidates and VOI is shown in Figure 5.1d. This ability to use VOI can thus facilitate the detection and tracking of markerballs.

Similar proof of concept trials to segment inflated balloon with min tree based filtering and VOI have also been part of the preliminary work. Though these methods need to be evaluated and validated, the above given example assures promising perspective directions of the thesis work. The following steps would include detection of the arrival and departure of a tool to further advance towards PCI procedure modeling. Several other conceivable applications of such modeling are mentioned in Section 1.3.

Digital stent enhancement technique, a temporal filtering technique, is used to produce an enhanced image of a stent from an X-ray image sequence. It is a special case of motion compensated noise reduction where motion of stent is inferred by motion of balloon markerballs [Bismuth et al., 2011; Koolen et al., 2005]. In these techniques, the detected VOI can help to augment the robustness of markerballs detection. Besides, the detected VOI can also be used as an aid to produce enhanced image of the stent in the absence of markerballs.

5.2.2 Cathlab of the future and Robotics in PCI

Today, the adoption of robotic assisted PCI systems has its fair share of supporters and critics. Operators already perform procedures manually without robotic assistance, but advocates for the technology point out these systems can greatly improve staff safety by reducing radiation dose and improve outcomes with more accurate procedural navigation [Carrozza, 2012; Granada et al., 2011; Bao et al., 2018]. There are now robotic systems available for peripheral, coronary and electrophysiology procedures. Users say these systems greatly improve their accuracy in vessel navigation and procedural precision, even among experienced well-known operators. These systems enable precise steering of guidewires and catheters through complex anatomies. The manufacturers of the robotic-assisted systems claim to make procedures less invasive, and to reduce physician's radiation dose by 95%. It is still important to reduce the patient's radiation and contrast agent dose. Enabling the imaging systems and the robotic navigation systems to recognize the phases of the procedure, through PCI procedure modelling, can help to reduce the patient dose. We hope that this work will establish a useful reference for further research in this exciting field.

Publication list

A.1 Journal

Bacchuwar, K., Cousty, J., Vaillant, R., and Najman, L. (2017). Scale-space for empty catheter segmentation in PCI fluoroscopic images. *International Journal of Computer Assisted Radiology and Surgery*, 12(7), 1179-1188.

A.2 Conferences

Bacchuwar, K., Cousty, J., Vaillant, R., and Najman, L. (September 2016). Towards semantic image analysis of PCI: Empty catheter segmentation. *CVII-Stent Workshop, MICCAI 2016*.

Bacchuwar, K., Cousty, J., Vaillant, R., and Najman, L. (2017). Scale-space for empty catheter segmentation in PCI fluoroscopic images. *IPCAI 2017, CARS*.

Bacchuwar, K., Cousty, J., Vaillant, R., and Najman, L. (2017). VOIDD: automatic vessel-of-intervention dynamic detection in PCI procedures. In *Intravascular Imaging and Computer Assisted Stenting, and Large-Scale Annotation of Biomedical Data and Expert Label Synthesis* (pp. 47-56). Springer, Cham., *MICCAI 2017*.

A.3 Patents

Riddell, C., Vaillant, R., and Bacchuwar, K. (General Electric Co, 2017) "System and method for adjusting a radiation dose during imaging of an object within a subject." U.S. Patent Application 15/087,304.

Bibliography

Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P. and Süsstrunk, S. [2012], ‘Slic superpixels compared to state-of-the-art superpixel methods’, *IEEE transactions on pattern analysis and machine intelligence* **34**(11), 2274–2282. Cited on page [42](#).

Agarwal, S., Joshi, A., Finin, T., Yesha, Y. and Ganous, T. [2007], ‘A pervasive computing system for the operating room of the future’, *Mob. Netw. Appl.* **12**(2-3), 215–228. Cited on page [12](#).

Aho, A. V., Hopcroft, J. E. and Ullman, J. D. [1983], ‘Data structures and algorithms’. Cited on page [42](#).

Aksamentov, I., Twinanda, A. P., Mutter, D., Marescaux, J. and Padoy, N. [2017], Deep neural networks predict remaining surgery duration from cholecystectomy videos, in ‘International Conference on Medical Image Computing and Computer-Assisted Intervention’, Springer, pp. 586–593. Cited on page [12](#).

Alt, H. and Godau, M. [1992], Measuring the resemblance of polygonal curves, in ‘Proceedings of the Eighth Annual Symposium on Computational Geometry’, SCG ’92, ACM, New York, NY, USA, pp. 102–109. Cited on page [92](#).

Ambrosini, P., Ruijters, D., Niessen, W. J., Moelker, A. and van Walsum, T. [2017], Fully automatic and real-time catheter segmentation in x-ray fluoroscopy, in ‘International Conference on Medical Image Computing and Computer-Assisted Intervention’, Springer, pp. 577–585. Cited on pages [ix](#) and [54](#).

Bacchuwar, K., Cousty, J., Vaillant, R. and Najman, L. [2016], ‘Towards semantic image analysis of PCI: Empty catheter segmentation’, *Proceeding of the CVII-STENT workshop at MICCAI 2016* . Cited on page [13](#).

Bacchuwar, K., Cousty, J., Vaillant, R. and Najman, L. [2017a], ‘Scale-space for empty catheter segmentation in PCI fluoroscopic images’, *International Journal*

of *Computer Assisted Radiology and Surgery* **12**(7), 1179–1188. Cited on pages [vii](#) and [13](#).

Bacchuwar, K., Cousty, J., Vaillant, R. and Najman, L. [2017b], Voidd: Automatic vessel-of-intervention dynamic detection in pci procedures, in ‘Intravascular Imaging and Computer Assisted Stenting, and Large-Scale Annotation of Biomedical Data and Expert Label Synthesis’, Springer International Publishing, pp. 47–56. Cited on pages [vii](#), [xii](#), and [14](#).

Baert, S. A., Viergever, M. A. and Niessen, W. J. [2003], ‘Guide-wire tracking during endovascular interventions’, *IEEE Transactions on Medical Imaging* **22**(8), 965–972. Cited on pages [16](#) and [54](#).

Ballester, C., Caselies, V. and Monasse, P. [2001], ‘The tree of shapes of an image’. Cited on page [43](#).

Bao, X., Guo, S., Xiao, N., Li, Y., Yang, C. and Jiang, Y. [2018], ‘A cooperation of catheters and guidewires-based novel remote-controlled vascular interventional robot’, *Biomedical microdevices* **20**(1), 20. Cited on page [115](#).

Barbu, A., Athitsos, V., Georgescu, B., Boehm, S., Durlak, P. and Comaniciu, D. [2007], Hierarchical learning of curves application to guidewire localization in fluoroscopy, in ‘2007 IEEE Conference on Computer Vision and Pattern Recognition’, IEEE, pp. 1–8. Cited on pages [16](#) and [57](#).

Baur, C., Albarqouni, S., Demirci, S., Navab, N. and Fallavollita, P. [2016], *Cath-Nets: Detection and Single-View Depth Prediction of Catheter Electrodes*, Springer International Publishing, pp. 38–49. Cited on pages [ix](#) and [54](#).

Benseghir, T. [2015], Topology Preserving Vascular Registration: Application to Percutaneous Coronary Intervention, Theses, Université de Nice-Sophia Antipolis. **URL:** <https://hal.inria.fr/tel-01235141> Cited on page [89](#).

Benseghir, T., Malandain, G. and Vaillant, R. [2015], ‘A tree-topology preserving pairing for 3d/2d registration’, *International Journal of Computer Assisted Radiology and Surgery* **10**(6), 913–923. Cited on pages [xiii](#), [50](#), [80](#), [92](#), and [93](#).

Berger, C., Géraud, T., Levillain, R., Widynski, N., Baillard, A. and Bertin, E. [2007], Effective component tree computation with application to pattern recognition in astronomical imaging, in ‘Image Processing, 2007. ICIP 2007. IEEE International Conference on’, Vol. 4, IEEE, pp. IV–41. Cited on pages [38](#) and [42](#).

Bertrand, G. [2005], ‘On topological watersheds’, *Journal of Mathematical Imaging and Vision* **22**(2-3), 217–230. Cited on pages [19](#) and [28](#).

Bertrand, G. and Couprie, M. [2007], Transformations topologiques discrètes, in C. J.-M. Coeurjolly David, Montanvert Annick, ed., ‘Géométrie discrète et images numériques’, Hermès, pp. 187–209. incollection. **URL:** <https://hal-upec-upem.archives-ouvertes.fr/hal-00622392> Cited on page [97](#).

Besl, P. J. and McKay, N. D. [1992], Method for registration of 3-d shapes, in ‘Sensor Fusion IV: Control Paradigms and Data Structures’, Vol. 1611, International Society for Optics and Photonics, pp. 586–607. Cited on page 49.

Bismuth, V., Vaillant, R., Funck, F., Guillard, N. and Najman, L. [2011], ‘A comprehensive study of stent visualization enhancement in X-ray images by image processing means’, *Medical Image Analysis* **15**(4), 565 – 576. Special section on IPMI 2009.

URL: <http://www.sciencedirect.com/science/article/pii/S1361841511000442>
Cited on pages 16 and 115.

Bismuth, V., Vaillant, R., Talbot, H. and Najman, L. [2012], Curvilinear structure enhancement with the polygonal path image-application to guide-wire segmentation in X-ray fluoroscopy, in ‘International Conference on Medical Image Computing and Computer-Assisted Intervention’, Springer, pp. 9–16. Cited on page 17.

Bismuth, V., Vancamberg, L. and Gorges, S. [2009], A comparison of line enhancement techniques-applications to guide-wire detection and respiratory motion tracking, in ‘Proc. of SPIE Vol’, Vol. 7259, pp. 72591M–1. Cited on page 16.

Bosilj, P., Kijak, E. and Lefèvre, S. [2018], ‘Partition and Inclusion Hierarchies of Images: A Comprehensive Survey’, *Journal of Imaging* **4**(2), 33. Cited on page 36.

Braga-Neto, U. and Goutsias, J. [2002], ‘Connectivity on complete lattices: New results’, *Computer Vision and Image Understanding* **85**(1), 22–53. Cited on page 47.

Brasselet, C., Addad, F. and Lafont, A. [2002], ‘Qu’est-ce qu’une sténose coronaire significative?’, *La Lettre du Cardiologue* (352), 32–36.

URL: <http://www.refdoc.fr/Detailnotice?idarticle=9908326> Cited on page 2.

Breen, E. J. and Jones, R. [1996], ‘Attribute openings, thinnings, and granulometries’, *Computer Vision and Image Understanding* **64**(3), 377–389. Cited on pages 25, 27, 44, and 45.

Brost, A., Liao, R., Hornegger, J. and Strobel, N. [2009], 3-d respiratory motion compensation during ep procedures by image-based 3-d lasso catheter model generation and tracking, in ‘International Conference on Medical Image Computing and Computer-Assisted Intervention’, Springer, pp. 394–401. Cited on page 54.

Brost, A., Liao, R., Strobel, N. and Hornegger, J. [2010], ‘Respiratory motion compensation by model-based catheter tracking during ep procedures’, *Medical Image Analysis* **14**(5), 695–706. Cited on page 54.

Caldairou, B., Naegel, B. and Passat, N. [2009], Segmentation of complex images based on component-trees: Methodological tools., in ‘ISMM’, Springer, pp. 171–180. Cited on page 19.

- Carlinet, E. and Géraud, T. [2014], ‘A comparative review of component tree computation algorithms’, *IEEE Transactions on Image Processing* **23**(9), 3885–3895. Cited on page [42](#).
- Carlinet, E. and Géraud, T. [2015], ‘MToS: A tree of shapes for multivariate images’, *IEEE Transactions on Image Processing* **24**(12), 5330–5342. Cited on page [36](#).
- Carrozza, J. P. [2012], ‘Robotic-assisted percutaneous coronary intervention—filling an unmet need’, *Journal of Cardiovascular Translational Research* **5**(1), 62–66. Cited on page [115](#).
- Caselles, V., Coll, B. and Morel, J.-M. [1999], ‘Topographic maps and local contrast changes in natural images’, *International Journal of Computer Vision* **33**(1), 5–27. Cited on page [39](#).
- Caselles, V. and Monasse, P. [2009], *Geometric description of images as topographic maps*, Springer. Cited on page [43](#).
- Chang, P. L., Rolls, A., Praetere, H. D., Poorten, E. V., Riga, C. V., Bicknell, C. D. and Stoyanov, D. [2016], ‘Robust catheter and guidewire tracking using b-spline tube model and pixel-wise posteriors’, *IEEE Robotics and Automation Letters* **1**(1), 303–308. Cited on page [54](#).
- Chen, B. J., Wu, Z., Sun, S., Zhang, D. and Chen, T. [2016], Guidewire tracking using a novel sequential segment optimization method in interventional X-ray videos, in ‘2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)’, pp. 103–106. Cited on pages [54](#), [80](#), and [89](#).
- Chen, L., Berry, M. W. and Hargrove, W. W. [2000], ‘Using dendronal signatures for feature extraction and retrieval’, *International Journal of Imaging Systems and Technology* **11**(4), 243–253. Cited on page [18](#).
- Chen, T., Wang, Y., Durlak, P. and Comaniciu, D. [2012], ‘Real time assistance for stent positioning and assessment by self-initialized tracking’, *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2012* pp. 405–413. Cited on page [16](#).
- Couprie, C., Grady, L., Najman, L. and Talbot, H. [2011], ‘Power watershed: A unifying graph-based optimization framework’, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **33**(7), 1384–1399. Cited on page [20](#).
- Couprie, M. and Bertrand, G. [2012], *Discrete Topological Transformations for Image Processing*, Springer Netherlands, Dordrecht, pp. 73–107. Cited on pages [xiii](#) and [90](#).
- Cousty, J., Bertrand, G., Najman, L. and Couprie, M. [2009], ‘Watershed cuts: Minimum spanning forests and the drop of water principle’, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **31**(8), 1362–1374. Cited on page [19](#).

Cousty, J., Najman, L., Dias, F. and Serra, J. [2013], ‘Morphological filtering on graphs’, *Computer Vision and Image Understanding* **117**(4), 370–385. Cited on page [22](#).

Cousty, J., Najman, L., Kenmochi, Y. and Guimarães, S. [2015], New characterizations of minimum spanning trees and of saliency maps based on quasi-flat zones, *in* ‘International Symposium on Mathematical Morphology and Its Applications to Signal and Image Processing’, Springer, pp. 205–216. Cited on page [63](#).

Cousty, J., Najman, L., Kenmochi, Y. and Guimarães, S. [2017], ‘Hierarchical segmentations with graphs: Quasi-flat zones, minimum spanning trees, and saliency maps’, *Journal of Mathematical Imaging and Vision* . Cited on pages [34](#) and [40](#).

Cousty, J., Najman, L. and Perret, B. [2013], Constructive links between some morphological hierarchies on edge-weighted graphs, *in* ‘International Symposium on Mathematical Morphology and Its Applications to Signal and Image Processing’, Springer, pp. 86–97. Cited on pages [18](#), [36](#), and [43](#).

Dannenberg, L., Polzin, A., Bullens, R., Kelm, M. and Zeus, T. [2016], ‘On the road: First-in-man bifurcation percutaneous coronary intervention with the use of a dynamic coronary road map and stentboost live imaging system’, *International Journal of Cardiology* **215**, 7–8. Cited on page [78](#).

Didier, R., Magalhaes, M., Koifman, E., Leven, F., Castellant, P., Boschat, J., Jobic, Y., Kiramijyan, S., Nicol, P. and Gilard, M. [2016], ‘The utilisation of the cardiovascular automated radiation reduction X-ray system (CARS) in the cardiac catheterisation laboratory aids in the reduction of the patient radiation dose.’, *EuroIntervention: journal of EuroPCR in collaboration with the Working Group on Interventional Cardiology of the European Society of Cardiology* **12**(8), e948. Cited on page [7](#).

Dotter, C. T. and Judkins, M. P. [1964], ‘Transluminal treatment of arteriosclerotic obstruction: description of a new technic and a preliminary report of its application’, *Circulation* **30**(5), 654–670. Cited on page [6](#).

Eiter, T. and Mannila, H. [1994], Computing discrete Fréchet distance, Technical report. Cited on pages [xiii](#) and [92](#).

Felzenszwalb, P. F. and Huttenlocher, D. P. [2004], ‘Efficient graph-based image segmentation’, *International journal of computer vision* **59**(2), 167–181. Cited on pages [42](#) and [45](#).

Florent, R., Nosjean, L., Lelong, P. and Rongen, P. M. J. [2008], ‘Medical viewing system and method for enhancing structures in noisy images’. US Patent 7,415,169. Cited on page [16](#).

Frangi, A. F., Niessen, W. J., Vincken, K. L. and Viergever, M. A. [1998], Multi-scale vessel enhancement filtering, *in* ‘Medical Image Computing and Computer-Assisted Intervention — MICCAI’98: First International Conference Cambridge,

MA, USA, October 11–13, 1998 Proceedings’, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 130–137. Cited on pages 16, 56, 80, and 89.

Frimerman, A., Abergel, E., Blondheim, D. S., Shotan, A., Meisel, S., Shochat, M., Punjabi, P. and Roguin, A. [2016], ‘Novel Method for Real Time Co-Registration of IVUS and Coronary Angiography’, *Journal of Interventional Cardiology* **29**(2), 225–231. Cited on pages xi and 78.

Gaspard, P. [2017], *The History of Coronary Angioplasty*, Europa Digital and Publishing. Cited on page 6.

Gatica-Perez, D., Gu, C., Sun, M.-T. and Ruiz-Correa, S. [2001], ‘Extensive partition operators, gray-level connected operators, and region merging/classification segmentation algorithms: theoretical links’, *IEEE Transactions on Image Processing* **10**(9), 1332–1345. Cited on pages 18 and 25.

Géraud, T., Carlinet, E., Crozet, S. and Najman, L. [2013], A quasi-linear algorithm to compute the tree of shapes of nD images, in ‘International Symposium on Mathematical Morphology and Its Applications to Signal and Image Processing’, Springer, pp. 98–110. Cited on page 43.

Granada, J. F., Delgado, J. A., Uribe, M. P., Fernandez, A., Blanco, G., Leon, M. B. and Weisz, G. [2011], ‘First-in-human evaluation of a novel robotic-assisted coronary angioplasty system’, *JACC: Cardiovascular Interventions* **4**(4), 460–465. Cited on page 115.

Grossiord, E., Talbot, H., Passat, N., Meignan, M., Tervé, P. and Najman, L. [2015], Hierarchies and shape-space for PET image segmentation, in ‘12th International Symposium on Biomedical Imaging (ISBI), 2015 IEEE’, IEEE, pp. 1118–1121. Cited on page 46.

Guigues, L., Cocquerez, J. P. and Le Men, H. [2006], ‘Scale-sets image analysis’, *International Journal of Computer Vision* **68**(3), 289–317. Cited on pages 19, 45, 56, and 57.

Guimarães, S., Kenmochi, Y., Cousty, J., Patrocínio, Z. and Najman, L. [2017], ‘Hierarchizing graph-based image segmentation algorithms relying on region dissimilarity: the case of the Felzenszwalb-Huttenlocher method’, *Mathematical Morphology - Theory and Applications*.
URL: <https://hal-ujec-upem.archives-ouvertes.fr/hal-01342967> Cited on page 42.

Havel, J., Merciol, F. and Lefèvre, S. [2013], Efficient schemes for computing α -tree representations, in ‘International Symposium on Mathematical Morphology and Its Applications to Signal and Image Processing’, Springer, pp. 111–122. Cited on page 43.

Havel, J., Merciol, F. and Lefèvre, S. [2016], ‘Efficient tree construction for multi-scale image representation and processing’, *Journal of Real-Time Image Processing* pp. 1–18. Cited on page 43.

Hebsgaard, L., Nielsen, T. M., Tu, S., Krusell, L. R., Maeng, M., Veien, K. T., Raungaard, B., Terkelsen, C. J., Kaltoft, A., Reiber, J. H., Lassen, J. F., Christiansen, E. H. and Holm, N. R. [2015], ‘Co-registration of optical coherence tomography and X-ray angiography in percutaneous coronary intervention. The Does Optical Coherence Tomography Optimize Revascularization (DOCTOR) fusion study’, *International Journal of Cardiology* **182**(Supplement C), 272 – 278. Cited on page [78](#).

Heibel, H., Glocker, B., Groher, M., Pfister, M. and Navab, N. [2013], ‘Interventional tool tracking using discrete optimization’, *IEEE transactions on medical imaging* **32**(3), 544–555. Cited on page [54](#).

Heibela, T. H., Glockera, B., Grohera, M., Paragios, N., Komodakis, N. and Navaba, N. [2009], Discrete tracking of parametrized curves, *in* ‘Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on’, IEEE, pp. 1754–1761. Cited on page [16](#).

Heijmans, H. J. [1991], ‘Theoretical aspects of gray-level morphology’, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **13**(6), 568–582. Cited on page [28](#).

Heijmans, H. J. and Vincent, L. [1999], ‘Graph morphology in image analysis’, *Mathematical Morphology in Image Processing*. Cited on page [20](#).

Hoffmann, M., Müller, S., Kurzydum, K., Strobel, N. and Hornegger, J. [2015], *Robust Identification of Contrasted Frames in Fluoroscopic Images*, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 23–28. Cited on page [81](#).

Honorat, N., Vaillant, R. and Paragios, N. [2010a], ‘Guide-wire extraction through perceptual organization of local segments in fluoroscopic images’, *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2010* pp. 440–448. Cited on page [17](#).

Honorat, N., Vaillant, R. and Paragios, N. [2010b], Robust guidewire segmentation through boosting, clustering and linear programming, *in* ‘Biomedical Imaging: From Nano to Macro, 2010 IEEE International Symposium on’, IEEE, pp. 924–927. Cited on page [17](#).

Honorat, N., Vaillant, R. and Paragios, N. [2012], Graph-based guide-wire segmentation through fusion of contrast-enhanced and fluoroscopic images, *in* ‘2012 9th IEEE International Symposium on Biomedical Imaging (ISBI)’, pp. 948–951. Cited on pages [80](#) and [89](#).

Jacob, M. and Unser, M. [2004], ‘Design of steerable filters for feature detection using canny-like criteria’, *IEEE transactions on pattern analysis and machine intelligence* **26**(8), 1007–1019. Cited on page [16](#).

- Jalba, A. C., Wilkinson, M. H. and Roerdink, J. B. [2006], ‘Shape representation and recognition through morphological curvature scale spaces’, *IEEE Transactions on Image Processing* **15**(2), 331–341. Cited on page [38](#).
- Jones, R. [1999], ‘Connected filtering and segmentation using component trees’, *Computer Vision and Image Understanding* **75**(3), 215–228. Cited on pages [18](#), [19](#), [25](#), and [38](#).
- Kim, D., Park, S., Jeong, M. H. and Ryu, J. [2017], ‘Registration of angiographic image on real-time fluoroscopic image for image-guided percutaneous coronary intervention’, *International Journal of Computer Assisted Radiology and Surgery* . Cited on page [49](#).
- Kirbas, C. and Quek, F. [2004], ‘A Review of Vessel Extraction Techniques and Algorithms’, *ACM Comput. Surv.* **36**(2), 81–121. Cited on page [80](#).
- Kita, Y., Wilson, D. L. and Noble, J. A. [1998], Real-time registration of 3D cerebral vessels to X-ray angiograms, in ‘International Conference on Medical Image Computing and Computer-Assisted Intervention’, Springer, pp. 1125–1133. Cited on page [50](#).
- Koenderink, J. J. [1984], ‘The structure of images’, *Biological cybernetics* **50**(5), 363–370. Cited on page [56](#).
- Koolen, J., Veer, M. and Hanekamp, C. [2005], ‘Stentboost image enhancement: first clinical experience’, *Kontraste (Hamburg)* **49**(2), 4–8. Cited on page [115](#).
- Koyama, K., Maehara, A., Fall, K. N., Mintz, G. S. and Ali, Z. A. [2015], ‘A case of on-line software-based co-registration of optical coherence tomography and angiography guided percutaneous coronary intervention for a patient with angina pectoris’, *International Journal of Cardiology* **201**, 484–486. Cited on pages [xi](#) and [78](#).
- Krissian, K., Malandain, G., Ayache, N., Vaillant, R. and Troussset, Y. [2000], ‘Model-based detection of tubular structures in 3d images’, *Computer Vision and Image Understanding* **80**(2), 130 – 171. Cited on pages [xiii](#) and [89](#).
- Kruskal, J. B. [1956], ‘On the shortest spanning subtree of a graph and the traveling salesman problem’, *Proceedings of the American Mathematical society* **7**(1), 48–50. Cited on page [43](#).
- Kunz, D. and Schweiger, B. [2005], ‘Line detection in strongly noise-corrupted images’, *Bildverarbeitung für die Medizin 2005* pp. 50–54. Cited on page [16](#).
- Lalys, F. and Jannin, P. [2014], ‘Surgical process modelling: a review’, *International Journal of Computer Assisted Radiology and Surgery* **9**(3), 495–511. Cited on pages [vii](#) and [11](#).

Lalys, F., Riffaud, L., Bouget, D. and Jannin, P. [2012], ‘A framework for the recognition of high-level surgical tasks from video images for cataract surgeries’, *IEEE TBME* **59**(4), 966–976. Cited on page [11](#).

Lessard, S., Kauffmann, C., Pfister, M., Cloutier, G., Thérasse, É., de Guise, J. A. and Soulez, G. [2015], ‘Automatic detection of selective arterial devices for advanced visualization during abdominal aortic aneurysm endovascular repair’, *Medical engineering & physics* **37**(10), 979–986. Cited on pages [ix](#) and [54](#).

Lézoray, O. and Grady, L. [2012], *Image processing and analysis with graphs: theory and practice*, CRC Press. Cited on page [19](#).

Lindeberg, T. [1993], ‘Detecting salient blob-like image structures and their scales with a scale-space primal sketch: a method for focus-of-attention’, *IJCV* **11**(3), 283–318. Cited on pages [16](#) and [56](#).

Liu, A., Bullitt, E. and Pizer, S. M. [1998], 3d/2d registration via skeletal near projective invariance in tubular objects, *in* ‘International Conference on Medical Image Computing and Computer-Assisted Intervention’, Springer, pp. 952–963. Cited on page [50](#).

Loménie, N. and Stamon, G. [2008], ‘Morphological mesh filtering and α -objects’, *Pattern Recognition Letters* **29**(10), 1571–1579. Cited on page [20](#).

Ma, Y., Gogin, N., Cathier, P., Housden, R. J., Gijbsbers, G., Cooklin, M., O’Neill, M., Gill, J., Rinaldi, C. A., Razavi, R. et al. [2013], ‘Real-time X-ray fluoroscopy-based catheter detection and tracking for cardiac electrophysiology interventions’, *Medical physics* **40**(7). Cited on page [54](#).

Ma, Y., King, A. P., Gogin, N., Rinaldi, C. A., Gill, J., Razavi, R. and Rhode, K. S. [2010], Real-time respiratory motion correction for cardiac electrophysiology procedures using image-based coronary sinus catheter tracking, *in* ‘International Conference on Medical Image Computing and Computer-Assisted Intervention’, Springer, pp. 391–399. Cited on pages [53](#) and [54](#).

Maintz, J. A. and Viergever, M. A. [1998], ‘A survey of medical image registration’, *Medical image analysis* **2**(1), 1–36. Cited on page [48](#).

Maragos, P. and Schafer, R. [1987], ‘Morphological filters—Part I: Their set-theoretic analysis and relations to linear shift-invariant filters’, *IEEE Transactions on Acoustics, Speech, and Signal Processing* **35**(8), 1153–1169. Cited on page [28](#).

Matas, P., Dokladalova, E., Akil, M., Grandpierre, T., Najman, L., Poupa, M. and Georgiev, V. [2008], Parallel algorithm for concurrent computation of connected component tree, *in* ‘International Conference on Advanced Concepts for Intelligent Vision Systems’, Springer, pp. 230–241. Cited on page [42](#).

Mattes, J., Richard, M. and Demongeot, J. [1999], Tree representation for image matching and object recognition, *in* ‘Discrete Geometry for Computer Imagery’, Springer, pp. 298–309. Cited on page [25](#).

- McLaughlin, R. A., Hipwell, J., Penney, G. P., Rhode, K., Chung, A., Noble, J. A. and Hawkes, D. [2001], ‘Intensity-based registration versus feature-based registration for neurointerventions’, *Proceedings of Medical Image Understanding Analysis*. Cited on page [49](#).
- Menotti, D., Najman, L. and de Albuquerque Araújo, A. [2007], 1D Component tree in linear time and space and its application to gray-level image multithresholding, *in* ‘Proceedings of the 8th International Symposium on Mathematical Morphology’, Citeseer, pp. 10–13. Cited on page [42](#).
- Meyer, F. and Angulo, J. [2007], ‘Micro-viscous morphological operators’, *Mathematical Morphology and its Application to Signal and Image Processing (ISMM 2007)* pp. 165–176. Cited on page [20](#).
- Meyer, F. and Lerallut, R. [2007], Morphological operators for flooding, leveling and filtering images using graphs, *in* ‘International Workshop on Graph-Based Representations in Pattern Recognition’, Springer, pp. 158–167. Cited on page [20](#).
- Meyer, F. and Maragos, P. [2000], ‘Nonlinear scale-space representation with morphological levelings’, *Journal of Visual Communication and Image Representation* **11**(2), 245–265. Cited on pages [ix](#), [36](#), and [40](#).
- Milletari, F., Belagiannis, V., Navab, N. and Fallavollita, P. [2014], *Fully Automatic Catheter Localization in C-Arm Images Using ℓ_1 -Sparse Coding*, Springer International Publishing, Cham, pp. 570–577. Cited on page [54](#).
- Milletari, F., Navab, N. and Fallavollita, P. [2013], Automatic detection of multiple and overlapping ep catheters in fluoroscopic sequences, *in* ‘International Conference on Medical Image Computing and Computer-Assisted Intervention’, Springer, pp. 371–379. Cited on page [53](#).
- Monasse, P. and Guichard, F. [2000*a*], ‘Fast computation of a contrast-invariant image representation’, *IEEE Transactions on Image Processing* **9**(5), 860–872. Cited on page [43](#).
- Monasse, P. and Guichard, F. [2000*b*], ‘Scale-space from a level lines tree’, *Journal of Visual Communication and Image Representation* **11**(2), 224–236. Cited on pages [18](#) and [43](#).
- Montero, R. S. and Bribiesca, E. [2009], State of the art of compactness and circularity measures, *in* ‘International mathematical forum’, Vol. 4, pp. 1305–1335. Cited on page [44](#).
- Morard, V., Decenciere, E. and Dokládál, P. [2013], ‘Efficient geodesic attribute thinnings based on the barycentric diameter’, *Journal of Mathematical Imaging and Vision* **46**(1), 128–142. Cited on page [44](#).
- Morard, V., Dokládál, P. and Decenciere, E. [2012], ‘One-Dimensional Openings, Granulometries and Component Trees in $O(1)$ Per Pixel’, *IEEE Journal of selected topics in signal processing* **6**(7), 840–848. Cited on page [42](#).

Nagao, M., Matsuyama, T. and Ikeda, Y. [1979], ‘Region extraction and shape analysis in aerial photographs’, *Computer Graphics and Image Processing* **10**(3), 195–223. Cited on page [40](#).

Najman, L. [2011], ‘On the equivalence between hierarchical segmentations and ultrametric watersheds’, *Journal of Mathematical Imaging and Vision* **40**(3), 231–247. Cited on pages [43](#) and [45](#).

Najman, L. and Couprie, M. [2006], ‘Building the component tree in quasi-linear time’, *IEEE Transactions on Image Processing* **15**(11), 3531–3539. Cited on pages [36](#) and [42](#).

Najman, L. and Cousty, J. [2014], ‘A graph-based mathematical morphology reader’, *Pattern Recognition Letters* **47**, 3–17. Cited on pages [20](#) and [57](#).

Najman, L., Cousty, J. and Perret, B. [2013], Playing with Kruskal: algorithms for morphological trees in edge-weighted graphs, in ‘International Symposium on Mathematical Morphology and Its Applications to Signal and Image Processing’, Springer, pp. 135–146. Cited on page [43](#).

Nallamothu, B. K., Tommaso, C. L., Anderson, H. V., Anderson, J. L., Dudley, R. A., Duffy, P. L., Faxon, D. P., Gurm, H. S., Hamilton, L. A., Jensen, N. C. et al. [2014], ‘ACC/AHA/SCAI/AMA–convened PCPI/NCQA 2013 performance measures for adults undergoing percutaneous coronary intervention’, *Journal of the American College of Cardiology* **63**(7), 722–745. Cited on pages [vii](#) and [12](#).

National Heart, Lung and Blood Institute [2015a], ‘Cardiac catheterization’.
URL: <https://www.nhlbi.nih.gov/health-topics/cardiac-catheterization> Cited on pages [8](#) and [9](#).

National Heart, Lung and Blood Institute [2015b], ‘Coronary artery bypass grafting (CABG)’.
URL: <https://www.nhlbi.nih.gov/health-topics/coronary-artery-bypass-grafting>
Cited on pages [xxiii](#) and [5](#).

National Heart, Lung and Blood Institute [2015c], ‘How is percutaneous coronary intervention done?’.
URL: <https://www.nhlbi.nih.gov/node/37777> Cited on pages [xxiii](#) and [5](#).

National Heart, Lung and Blood Institute [2015d], ‘Plaque buildup in an artery’.
URL: <https://www.nhlbi.nih.gov/health-topics/atherosclerosis> Cited on pages [xxiii](#) and [3](#).

Nistér, D. and Stewénus, H. [2008], ‘Linear time maximally stable extremal regions’, *Computer Vision–ECCV 2008* pp. 183–196. Cited on page [42](#).

OECD [2011], ‘Cardiac procedures (coronary angioplasty)’.
URL: <https://goo.gl/PmQr0> Cited on pages [vi](#) and [11](#).

- Oliveira, F. P. and Tavares, J. M. R. [2014], ‘Medical image registration: a review’, *Computer methods in biomechanics and biomedical engineering* **17**(2), 73–93. Cited on page 48.
- Ouzounis, G. K. and Soille, P. [2011], Pattern spectra from partition pyramids and hierarchies., *in* ‘ISMM’, Springer, pp. 108–119. Cited on page 40.
- Ouzounis, G. K. and Wilkinson, M. H. [2007a], ‘Mask-based second-generation connectivity and attribute filters’, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **29**(6), 990–1004. Cited on pages 19 and 47.
- Ouzounis, G. K. and Wilkinson, M. H. [2007b], A parallel implementation of the dual-input max-tree algorithm for attribute filtering, *in* ‘Proc. 8th Int. Symp. Math. Morphol.(ISMM)’, Vol. 1, pp. 449–460. Cited on page 42.
- Ouzounis, G. K. and Wilkinson, M. H. [2011], ‘Hyperconnected attribute filters based on k-flat zones’, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **33**(2), 224–239. Cited on page 19.
- Padoy, N., Blum, T., Ahmadi, S.-A., Feussner, H., Berger, M.-O. and Navab, N. [2012], ‘Statistical modeling and recognition of surgical workflow’, *Medical image analysis* **16**(3), 632–641. Cited on page 11.
- Palti-Wasserman, D., Brukstein, A. M. and Beyar, R. P. [1997], ‘Identifying and tracking a guide wire in the coronary arteries during angioplasty from X-ray images’, *IEEE Transactions on Biomedical Engineering* **44**(2), 152–164. Cited on page 16.
- Passat, N. and Naegel, B. [2011], Component-hypertrees for image segmentation, *in* ‘International Symposium on Mathematical Morphology and Its Applications to Signal and Image Processing’, Springer, pp. 284–295. Cited on page 19.
- Passat, N., Naegel, B., Rousseau, F., Koob, M. and Dietemann, J.-L. [2011], ‘Interactive segmentation based on component-trees’, *Pattern Recognition* **44**(10), 2539–2554. Cited on pages 18 and 25.
- Perret, B. and Collet, C. [2015], ‘Connected image processing with multivariate attributes: An unsupervised markovian classification approach’, *Computer Vision and Image Understanding* **133**, 1–14. Cited on page 45.
- Perret, B., Cousty, J., Tankyevych, O., Talbot, H. and Passat, N. [2015], ‘Directed connected operators: asymmetric hierarchies for image filtering and segmentation’, *IEEE transactions on pattern analysis and machine intelligence* **37**(6), 1162–1176. Cited on pages 19, 23, 29, and 30.
- Postolski, M., Couprie, M. and Janaszewski, M. [2013], *Scale Filtered Euclidean Medial Axis*, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 360–371. Cited on page 97.

Prasad, M., Cassar, A., Fetterly, K. A., Bell, M., Theessen, H., Ecabert, O., Bresnahan, J. F. and Lerman, A. [2016], ‘Co-registration of angiography and intravascular ultrasound images through image-based device tracking’, *Catheterization and Cardiovascular Interventions* **88**(7), 1077–1082. Cited on page [78](#).

Riddell, C., Vaillant, R. and Bacchuwar, K. S. [2017], ‘System and method for adjusting a radiation dose during imaging of an object within a subject’. US Patent App. 15/087,304. Cited on page [12](#).

Ronse, C. [1998], ‘Set-theoretical algebraic approaches to connectivity in continuous or digital spaces’, *Journal of Mathematical Imaging and Vision* **8**(1), 41–58. Cited on pages [19](#) and [47](#).

Ronse, C. [2006], ‘Flat morphology on power lattices’, *Journal of Mathematical Imaging and Vision* **26**(1-2), 185. Cited on page [28](#).

Ronse, C. [2014], ‘Ordering partial partitions for image segmentation and filtering: Merging, creating and inflating blocks’, *Journal of mathematical imaging and vision* **49**(1), 202–233. Cited on page [34](#).

Saha, P. K., Borgfors, G. and di Baja, G. S. [2016], ‘A survey on skeletonization algorithms and their applications’, *Pattern Recognition Letters* **76**(Supplement C), 3 – 12. Special Issue on Skeletonization and its Application. Cited on pages [90](#) and [92](#).

Salembier, P. and Garrido, L. [2000], ‘Binary partition tree as an efficient representation for image processing, segmentation, and information retrieval’, *IEEE transactions on Image Processing* **9**(4), 561–576. Cited on pages [18](#) and [48](#).

Salembier, P., Oliveras, A. and Garrido, L. [1998], ‘Antiextensive connected operators for image and sequence processing’, *IEEE Transactions on Image Processing* **7**(4), 555–570. Cited on pages [18](#), [23](#), [25](#), [27](#), [36](#), [37](#), [38](#), [42](#), [44](#), and [45](#).

Salembier, P. and Serra, J. [1995], ‘Flat zones filtering, connected operators, and filters by reconstruction’, *IEEE Transactions on image processing* **4**(8), 1153–1160. Cited on pages [vii](#), [18](#), [23](#), and [25](#).

Salembier, P. and Wilkinson, M. H. F. [2009], ‘Connected operators’, *IEEE Signal Processing Magazine* **26**(6), 136–157. Cited on pages [vii](#), [viii](#), [18](#), [23](#), [48](#), [56](#), [58](#), and [80](#).

Sanborn, T. A., Tcheng, J. E., Anderson, H. V., Chambers, C. E., Cheatham, S. L., DeCaro, M. V., Durack, J. C., Everett, A. D., Gordon, J. B., Hammond, W. E. et al. [2014], ‘ACC/AHA/SCAI 2014 health policy statement on structured reporting for the cardiac catheterization laboratory’, *Journal of the American College of Cardiology* **63**(23), 2591–2623. Cited on pages [vii](#) and [12](#).

Schoonenberg, G., Florent, R., Lelong, P., Wink, O., Ruijters, D., Carroll, J. and ter Haar Romeny, B. [2009], ‘Projection-based motion compensation and reconstruction of coronary segments and cardiac implantable devices using rotational X-ray angiography’, *Medical image analysis* **13**(5), 785–792. Cited on page 16.

Schoonenberg, G., Lelong, P., Florent, R., Wink, O. and ter Haar Romeny, B. [2008], *The Effect of Automated Marker Detection on in Vivo Volumetric Stent Reconstruction*, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 87–94. Cited on page 16.

Serra, J. [1982], ‘Image analysis and mathematical morphology. vol i, and image analysis and mathematical morphology. vol ii: Theoretical advances’. Cited on pages 25 and 28.

Serra, J. [1988], *Image analysis and mathematical morphology: Theoretical advances*, Image Analysis and Mathematical Morphology, Academic Press.
URL: <https://books.google.fr/books?id=BpdTAAAYAAJ> Cited on pages 20, 23, and 24.

Serra, J. [2012], ‘Tutorial on connective morphology’, *IEEE Journal of Selected Topics in Signal Processing* **6**(7), 739–752. Cited on page 19.

Serra, J. C. and Salembier, P. [1993], Connected operators and pyramids, in ‘Image algebra and morphological image processing IV’, Vol. 2030, International Society for Optics and Photonics, pp. 65–77. Cited on pages vii, 18, 23, 24, and 25.

Shaked, D. and Bruckstein, A. M. [1998], ‘Pruning medial axes’, *Computer Vision and Image Understanding* **69**(2), 156 – 169. Cited on page 92.

Slabaugh, G., Kong, K., Unal, G. and Fang, T. [2007], ‘Variational guidewire tracking using phase congruency’, *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2007* pp. 612–619. Cited on page 16.

Soille, P. [2008], ‘Constrained connectivity for hierarchical image partitioning and simplification’, *PAMI, IEEE Transactions* **30**(7), 1132–1145. Cited on pages 19, 32, 40, and 63.

Soille, P. and Najman, L. [2012], On morphological hierarchical representations for image processing and spatial data clustering, in ‘Applications of Discrete Geometry and Mathematical Morphology’, Springer, pp. 43–67. Cited on page 43.

Stawiaski, J. and Meyer, F. [2009], Minimum spanning tree adaptive image filtering, in ‘Image Processing (ICIP), 2009 16th IEEE International Conference on’, IEEE, pp. 2245–2248. Cited on page 20.

Ta, V.-T., Elmoataz, A. and Lézoray, O. [2008], Partial difference equations over graphs: Morphological processing of arbitrary discrete data, in ‘European Conference on Computer Vision’, Springer, pp. 668–680. Cited on page 20.

Tarjan, R. E. [1975], ‘Efficiency of a good but not linear set union algorithm’, *Journal of the ACM (JACM)* **22**(2), 215–225. Cited on pages [42](#) and [43](#).

Tu, S., Holm, N. R., Koning, G., Huang, Z. and Reiber, J. H. C. [2011], ‘Fusion of 3D QCA and IVUS/OCT’, *The International Journal of Cardiovascular Imaging* **27**(2), 197–207. Cited on page [78](#).

Turaga, P., Chellappa, R., Subrahmanian, V. S. and Udrea, O. [2008], ‘Machine recognition of human activities: A survey’, *IEEE Transactions on Circuits and Systems for Video technology* **18**(11), 1473–1488. Cited on pages [vii](#) and [11](#).

Urbach, E., Boersma, N. and Wilkinson, M. [2005], ‘Vector-attribute filters’, *Mathematical Morphology: 40 Years On* pp. 95–104. Cited on page [45](#).

Urbach, E. R., Roerdink, J. B. and Wilkinson, M. H. [2007], ‘Connected shape-size pattern spectra for rotation and scale-invariant classification of gray-scale images’, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **29**(2), 272–285. Cited on pages [18](#), [25](#), [27](#), [44](#), and [45](#).

Vandini, A., Glocker, B., Hamady, M. and Yang, G.-Z. [2017], ‘Robust guidewire tracking under large deformations combining segment-like features (seglets)’, *Medical Image Analysis* **38**(Supplement C), 150 – 164. Cited on pages [80](#) and [89](#).

Vincent, L. [1989], ‘Graphs and mathematical morphology’, *Signal Processing* **16**(4), 365–388. Cited on page [19](#).

Wagner, M. G., Strother, C. M. and Mistretta, C. A. [2016], Guidewire path tracking and segmentation in 2D fluoroscopic time series using device paths from previous frames, *in* ‘SPIE Medical Imaging’, International Society for Optics and Photonics, pp. 97842B–97842B. Cited on page [54](#).

Wang, L., Xie, X. L., Bian, G. B., Hou, Z. G., Cheng, X. R. and Prasong, P. [2017], Guide-wire detection using region proposal network for X-ray image-guided navigation, *in* ‘2017 International Joint Conference on Neural Networks (IJCNN)’, pp. 3169–3175. Cited on pages [80](#) and [89](#).

Wang, P., Chen, T., Zhu, Y., Zhang, W., Zhou, S. K. and Comaniciu, D. [2009], Robust guidewire tracking in fluoroscopy, *in* ‘2009 IEEE Conference on Computer Vision and Pattern Recognition’, pp. 691–698. Cited on pages [54](#), [80](#), and [89](#).

Wang, P., Ecabert, O., Chen, T., Wels, M., Rieber, J., Ostermeier, M. and Comaniciu, D. [2013], ‘Image-based co-registration of angiography and intravascular ultrasound images’, *IEEE Transactions on Medical Imaging* **32**(12), 2238–2249. Cited on page [78](#).

Wang, Y., Chen, T., Wang, P., Rohkohl, C. and Comaniciu, D. [2012], ‘Automatic localization of balloon markers and guidewire in rotational fluoroscopy with application to 3d stent reconstruction’, *Computer Vision–ECCV 2012* pp. 428–441. Cited on page [16](#).

- Wendt, P., Coyle, E. and Gallagher, N. [1986], ‘Stack filters’, *IEEE Transactions on Acoustics, Speech, and Signal Processing* **34**(4), 898–911. Cited on page 28.
- Westenberg, M. A., Roerdink, J. B. and Wilkinson, M. H. [2007], ‘Volumetric attribute filtering and interactive visualization using the max-tree representation’, *IEEE Transactions on Image Processing* **16**(12), 2943–2952. Cited on page 38.
- Wiedemann, C., Heipke, C., Mayer, H. and Jamet, O. [1998], ‘Empirical evaluation of automatically extracted road axes’, *Empirical Evaluation Techniques in Computer Vision* pp. 172–187. Cited on page 66.
- Wilkinson, M. H. [2011], A fast component-tree algorithm for high dynamic-range images and second generation connectivity, in ‘Image Processing (ICIP), 2011 18th IEEE International Conference on’, IEEE, pp. 1021–1024. Cited on page 42.
- Wilkinson, M. H., Gao, H., Hesselink, W. H., Jonker, J.-E. and Meijster, A. [2008], ‘Concurrent computation of attribute filters on shared memory parallel machines’, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **30**(10), 1800–1813. Cited on page 42.
- Witkin, A. [1984], Scale-space filtering: A new approach to multi-scale description, in ‘Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP’84.’, Vol. 9, IEEE, pp. 150–153. Cited on page 56.
- World Health Organization [2017], ‘Top 10 causes of death worldwide’.
URL: <http://www.who.int/mediacentre/factsheets/fs310/en/> Cited on page 2.
- Wu, W., Chen, T., Wang, P., Zhou, S. K., Comaniciu, D., Barbu, A. and Strobel, N. [2011], Learning-based hypothesis fusion for robust catheter tracking in 2D X-ray fluoroscopy, in ‘Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on’, IEEE, pp. 1097–1104. Cited on pages ix and 53.
- Wu, X., Housden, J., Ma, Y., Razavi, B., Rhode, K. and Rueckert, D. [2015], ‘Fast catheter segmentation from echocardiographic sequences based on segmentation from corresponding X-ray fluoroscopy for cardiac catheterization interventions’, *IEEE transactions on medical imaging* **34**(4), 861–876. Cited on page 54.
- Xu, Y. [2013], Tree-based shape spaces: Definition and applications in image processing and computer vision, PhD thesis, Université Paris-Est, Marne-la-Vallée, France. Cited on pages 37, 45, and 46.
- Xu, Y., Géraud, T. and Najman, L. [2016], ‘Connected filtering on tree-based shape-spaces’, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **38**(6), 1126–1140. Cited on pages viii, xiii, 19, 45, 46, 57, and 90.
- Zhang, D. P. [2010], Coronary artery segmentation and motion modelling, Theses, Imperial College London.
URL: <http://hdl.handle.net/10044/1/6367> Cited on page 80.

Zitova, B. and Flusser, J. [2003], ‘Image registration methods: a survey’, *Image and vision computing* **21**(11), 977–1000. Cited on page [48](#).

- Adjacency relation, 20
- Angiography, 2
- Angioplasty, 5
- Anti-extensive, 23
- Attribute function, 24
 - area, 26
 - height, 26
 - volume, 26
- Balloon, 8
- Cathlab, 5
- Closing, 23
- Component, 20
 - of graph, 20
 - of stack, 28
- Component hierarchy, 34
- Component tree, 36
- Connected component, 20
- Connected component partition, 33
- Connected filters, 18
- Connection, 19
- Depth, 33
- Edge, 19
- Extensive, 23
- Flat zone, 24
- Fluoroscopy, 6
- Graph, 19
 - connected, 20
 - weighted, 28
- Guidewire, 8
- Hierarchy, 33
 - complete, 33
- Idempotent, 23
- Increasing, 23
- Leaf, 36
- Lower-level set, 28
- Markerballs, 8
- Max tree, 36
 - of image, 37
 - of weighted graph, 36
- Min tree, 36
 - of image, 36
 - of weighted graph, 36
- Node, 18, 35
- Opening, 23
- Operator, 22
 - binary, 22
 - connected, 23
 - grayscale, 22
- Partition, 31
 - connected, 33
 - partial, 33
- Path, 20
- PCI, 5
- Quasi-flat zones tree, 39

Record, [6](#)
Region, [33](#)
 of hierarchy, [35](#)
 of partition, [33](#)
Root, [36](#)

Self-dual, [23](#)
Stack, [28](#)
 Edge induced, [31](#)
 Vertex induced, [29](#)
Stent, [8](#)
StentViz, [9](#)
Subgraph, [20](#)

Tree of Shapes, [42](#)
Tree-based shape space, [36](#)

Union-find, [42](#)
Upper-level set, [28](#)

Vertex, [19](#)
Vessel of Intervention, [76](#)

Abstract: Percutaneous coronary interventions (PCI) are performed using real-time X-ray fluoroscopic images in an interventional suite. PCI procedure modeling enables the interventional machine to identify, monitor and understand the different phases of the procedure. Such modeling can help the practitioners by improving their interaction with the machine and by optimizing the X-ray dose and the contrast agent. One of the important tasks in achieving this goal is to segment different interventional tools into the flow of X-ray fluoroscopic images and to derive semantic information from them. The component tree, a powerful mathematical morphological tool, forms the basis of the proposed segmentation methods. We present this work in two parts: 1) empty catheter segmentation, and 2) vessel of intervention dynamic detection (VOIDD). We present a new multi-scale space-based segmentation method for detecting low-contrasted objects such as an empty catheter. For the last part, we present the segmentation of the guidewire tip with filtering based on the component tree and propose an algorithm to semantically track the segmented tip to determine the vessel of intervention. The qualitative and quantitative experimental results prove the efficiency of the proposed methods and their potential to aid the detection of interventional tools such as balloon markerballs, balloons and stents.

Résumé: Les interventions coronariennes percutanées (ICP) sont effectuées dans des salles interventionnelles à l'aide d'un guidage temps-réel d'images rayons X fluoroscopiques. La modélisation des procédures ICP permet d'identifier automatiquement et de comprendre ses différentes phases. Cette modélisation peut aider les praticiens à améliorer leurs interactions avec le système interventionnel et à optimiser la dose de rayon X et de produit de contraste. Pour atteindre cet objectif, l'une des tâches importantes consiste à segmenter différents outils de l'intervention dans le flux d'images fluoroscopiques et à en déduire des informations sémantiques. L'arbre des composantes, un puissant outil de morphologie mathématique, constitue la base des méthodes de segmentation proposées. Nous présentons ce travail en deux parties: 1) la segmentation du cathéter vide, et 2) la segmentation de la pointe du guide et le suivi de la détection du vaisseau d'intérêt. Nous présentons une nouvelle méthode de segmentation basée sur l'espace multi-échelles pour détecter des objets faiblement contrastés tels que le cathéter vide. Pour la dernière partie, nous présentons la segmentation de la pointe du guide avec un filtrage basé sur l'arbre des composantes et proposons un algorithme de suivi sémantique permettant de segmenter le vaisseau d'intérêt. Les résultats expérimentaux démontrent l'efficacité des méthodes proposées et leur potentiel pour la détection des outils interventionnels tels que les marqueurs de ballons, les ballons d'angioplastie et les stents.