



HAL
open science

Segmentation Spatio-temporelle d'une séquence d'images satellitaires à haute résolution.

Camille Le Men

► **To cite this version:**

Camille Le Men. Segmentation Spatio-temporelle d'une séquence d'images satellitaires à haute résolution.. Traitement des images [eess.IV]. Ecole nationale supérieure des telecommunications - ENST, 2009. Français. NNT: . pastel-00658159

HAL Id: pastel-00658159

<https://pastel.hal.science/pastel-00658159>

Submitted on 10 Jan 2012

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.



École Doctorale
d'Informatique,
Télécommunications
et Électronique de Paris

Thèse

présentée pour obtenir le grade de docteur
de l'Ecole Nationale Supérieure des Télécommunications

Spécialité : Signal et Images

Camille LE MEN

Segmentation spatio-temporelle d'une
séquence temporelle d'images satellitaires
à haute résolution.

Date prévue le 8 septembre 2009 devant le jury composé de

Philippe Bolon
Grégoire Mercier
Walter Kropatsch
Philippe Réfrégier
Alain Giros
Henri Maître
Mihai Datcu

Rapporteurs

Examineurs

Directeurs de thèse

Remerciements

Cette thèse s'est déroulée dans l'équipe du CoC (Center Of Competence on Information Extraction and Image Understanding for Earth Observation) née d'une collaboration entre le Centre National d'Etudes Spatiales (CNES), l'agence spatiale allemande (DLR) et Telecom Paristech.

Mes remerciements s'adressent en premier lieu à mon directeur, Henri Maître, et à mon co-directeur, Mihai Datcu, pour leur accompagnement tout au long de cette recherche ainsi que pour leurs fructueuses suggestions. Mihai Datcu, Marine Campedel et Alain Giros se sont particulièrement investis dans l'organisation de cette équipe, ce qui a donné lieu à de nombreux échanges lors de réunions et séminaires qui se sont tenus à Toulouse, Paris, et Munich. Des discussions enrichissantes avec Henri Maître m'ont permis d'orienter mes travaux sur la recherche de motifs fréquents dans les graphes.

Les cours de DEA de Philippe Réfrégier ont été à l'origine d'une réflexion sur l'application du principe de la minimisation de longueur de description à la segmentation jointe des STIS. Des suggestions de Mihai Datcu et de Lionel Gueguen les ont enrichies, ce qui a orienté ce mémoire vers le codage et la théorie de l'information.

Les échanges scientifiques sur la fouille de données avec Nicolas Méger et Andreea Julea, rencontrés lors d'une conférence, ont été particulièrement fructueux. Ils m'ont introduite à ce domaine de recherche, et notre collaboration a depuis débouché sur un article commun. Le développement de la thèse sur la recherche de motifs fréquents dans les graphes est né de ces discussions.

Je tiens à remercier particulièrement Alain Giros pour de fréquentes discussions tout au long de ma thèse, son soutien moral constant, et ses nombreuses relectures (jusque depuis l'aéroport d'Amsterdam lors d'un transit pour la Chine!). Merci à David pour ses conseils en programmation, ainsi qu'à Olivier qui m'a initiée à la bibliothèque ITK, à cmake (et m'a appris à réagir aux insultes de gdb!).

Merci par avance aux professeurs et experts qui ont accepté de participer à ce jury, Grégoire Mercier et Philippe Bolon, en tant que rapporteurs, Alain Giros, Walter Kropatsch et Philippe Réfrégier, en tant qu'examineurs. Ils ont eu la bienveillance de consacrer à la lecture de cette thèse terminée tardivement une partie de leurs vacances.

Je remercie aussi toute l'équipe du CoC, ainsi que les thésards et permanents de l'équipe TSI pour l'ambiance sympathique qui y règne. Enfin, je remercie Cyrielle, Jérémie et Olivier pour leur soutien quotidien, ainsi que David lors de ses séjours en France. Merci aussi à mes parents, et à mes frère et sœurs pour leur soutien, leurs encouragements, et leurs relectures!

Résumé

Les séquences temporelles d'images présentent une quantité d'information bien plus importante que des images individuelles. En effet, la prise en compte du temps accroît considérablement le nombre d'états possibles ce qui se traduit par une quantité d'information plus importante (au sens entropie).

Grâce à l'agilité des satellites de nouvelle génération et à leur utilisation dans des constellations, des séquences temporelles d'images satellitaires, STIS, vont maintenant être accessibles à haute résolution. Ces données sont extrêmement riches, mais en contrepartie, elles sont complexes et difficiles à interpréter manuellement. Des méthodes d'analyse automatique sont donc requises.

Les STIS à haute résolution (STISHR) se distinguent des autres types de séquences existantes par la nature des changements qu'elles présentent. En effet, contrairement aux STIS à basse résolution, les STISHR contiennent des objets. Afin de tenir compte de cette particularité, une méthode d'analyse orientée objet doit donc être utilisée. De telles méthodes existent dans le domaine de la vidéo. Cependant, les objets des STISHR peuvent subir des changements radiométriques, par exemple liés à la croissance des plantes, alors que la radiométrie des objets de vidéo est généralement supposée invariante. Nous devons donc concevoir une méthode d'analyse orientée objet adaptée à la STISHR. Par ailleurs, l'échantillonnage temporel des STISHR est irrégulier et généralement sous-échantillonné par rapport aux phénomènes observables à ces résolutions spatiales. De plus, du point de vue du satellite, aux évolutions radiométriques propres aux objets de la scène se rajoutent les évolutions radiométriques liées à l'épaisseur atmosphérique. Néanmoins, hormis de faibles erreurs de recalage géométrique, la STISHR présente une redondance temporelle dans la forme des objets. Ainsi, une construction est généralement pérenne et une zone cultivable est rarement modifiée. Nous proposons donc une méthode d'analyse de la STISHR basée objet de façon à exploiter d'une part la redondance radiométrique intra objets spatiaux, et d'autre part, la redondance temporelle de forme entre ces objets.

Une étude phénoménologique de la dynamique de la scène et de la STISHR nous permet d'identifier les caractéristiques d'une représentation adaptée à la description de ses changements : il s'agit d'un graphe dont les nœuds sont des objets spatiaux reliés par des arcs exprimant leur dépendance temporelle. Le calcul de ce graphe est cependant un problème difficile, et nous proposons d'en calculer une approximation. Cette approximation est calculée en deux temps. Dans un premier temps, on considère une approximation forte sur les dépendances temporelles afin de faciliter la détermination des régions spatiales. Celles-ci sont extraites grâce à un algorithme de segmentation exploitant conjointement les deux types de redondances disponibles : l'homogénéité radiométrique spatiale, et la redondance géométrique temporelle. Les nœuds du graphe ainsi déterminés, on relâche l'approximation sur les dépendances temporelles de façon à obtenir une estimation plus fine du graphe.

Nous proposons ensuite deux utilisations de ce graphe. L'une exploite ses caractéristiques structurelles telles que la valence des nœuds afin de détecter les changements géométriques. L'autre, utilise une version attribuée par des propriétés radiométriques du graphe pour mettre en évidence des comportements fréquents.

Abstract

Image Time series represent an amount of information far greater than individual images. In fact, considering time increases significantly the number of possible states corresponding to a bigger amount of information (in an entropic sense).

Thanks to the new generation satellites agility and to their use in constellations, satellite image time series, SITS, will from now on be available at higher resolution. These data are extremely rich, but, in compensation, are complex and difficult to interpret manually. Automatic analysis methods are thus required.

High resolution SITS (HRSITS) differ from the other types of existing sequences by the type of change they show. In fact, contrary to low resolution SITS, HRSITS present objects. In order to take this particularity into account, an object oriented analysis must be performed. Such approaches exist in the video domain. However, the HRSITS' objects may undergo changes in radiometry (for example, due to the plants growth) whereas the video object's radiometry is generally assumed time invariant. We thus need to design an object oriented analysis adapted to HRSITS. Besides, the temporal sampling of the HRSITS is irregular and generally sub-sampled compared to visible phenomena at the considered spatial resolution. Moreover, from the satellite viewpoint, some radiometry changes due to the atmosphere thickness cumulate with the radiometry evolution of the scene objects. Nevertheless, despite slight geometrical registration errors, the HRSITS' objects shape is temporally redundant. A building is in fact generally sustainable whereas an agricultural zone is rarely modified. We thus propose an object based HRSITS analysis method exploiting on one hand the intra-objects radiometry redundancy and on the other hand the temporal redundancy of the shape of the objects different versions.

A phenomenological study of the dynamic of the scene and of the HRSITS allows us to identify the characteristics of an adapted representation of the changes description : a graph the nodes of which are the spatial objects linked by edges representing their temporal dependency. However, the computation of this graph is a difficult problem, and we thus propose to compute an approximation in two steps. In the first step, a strong approximation is considered on the temporal dependencies in order to facilitate the determination of spatial regions. These are then extracted thanks to a segmentation algorithm exploiting jointly the two kind of available redundancies, namely the spatial homogeneity of the radiometry and the temporal geometric redundancy.

The nodes of the graph are thus determined and the approximation on the temporal dependencies can then be relaxed in order to achieve a finer estimation of the graph.

We then propose two utilisations of the graph. The first one exploits the structural characteristics such as the nodes degrees in order to detect and qualify geometric changes. The other uses an attributed version of the graph with radiometry properties attached to the nodes in order to extract relevant frequent behaviour.

Table des matières

Introduction	19
0.1 Intérêt des STIS	21
0.2 Position du problème	21
I Etat de l’art	27
1 Différentes séquences, différentes méthodes d’analyse et différents modèles	29
1.1 Position de notre problème par rapport à l’état de l’art	29
2 Modélisation de données spatio-temporelles	33
2.1 Modèle de base de données spatio-temporelles	33
2.1.1 Modélisation du temps	33
2.1.2 Modélisation de l’espace spatio-temporel	34
2.1.3 Modèle de base de données spatio-temporelles	35
2.2 Conclusion	38
II Présentation des données et du problème	39
3 Présentation des données	41
3.1 La STIS ADAM	41
3.1.1 Les images SPOT	41
3.1.2 Sélection des images	42
3.1.3 Prétraitements	42
3.1.4 Correction radiométrique	43
3.1.5 La scène observée : phénoménologie	44
3.2 Approche adoptée	54
III Cadre théorique	55
4 Approche de l’estimation du modèle de STIS	57
4.1 Scène statique	57
4.2 Vers un modèle de scène dynamique de télédétection à haute résolution spatiale et temporelle	57
4.2.1 Profondeur des objets	58
4.2.2 Définition d’un objet spatio-temporel (OST)	58

4.2.3	Processus des OST	58
4.2.4	Événements des OST	59
4.2.5	OST et sémantique	59
4.2.6	Conclusion	60
4.3	Hypothèses simplificatrices pour l'estimation d'un modèle de STIS	61
4.3.1	Complexité de l'estimation	61
4.3.2	Sous-échantillonnage temporel	61
4.3.3	Hypothèse de scène arrière plan	61
4.3.4	Conclusion	62
4.4	Approche de l'analyse de l'arrière plan	62
4.4.1	Description de la scène par un graphe	63
4.4.2	Présentation de la méthode d'estimation du modèle de STIS	64
4.5	Conclusion	66
5	Segmentation	67
5.1	Différentes approches pour la segmentation d'images	67
5.2	Segmentation d'images par minimisation de la longueur de description	69
5.2.1	Longueur de description	70
5.2.2	Optimisation	72
6	Recherche d'ensembles de symboles fréquents	77
6.1	Problème et définitions	77
6.2	Propriétés générales	78
6.3	ECLAT	80
IV	Mise en oeuvre et implémentation	83
7	Segmentation jointe d'une STIS	85
7.1	Introduction	85
7.2	Modèle géométrique spatio-temporel	86
7.2.1	Codage différentiel de la géométrie	86
7.2.2	Image de référence	87
7.2.3	Différents types de changements	87
7.2.4	Description idéale et complexité	87
7.2.5	Décision concernant le choix du type de changements	88
7.2.6	Décision concernant le choix de l'image de référence	89
7.3	Longueur de description du modèle géométrique spatio-temporel	90
7.3.1	Calcul du cylindre d'une région	90
7.3.2	Longueur de description géométrique conditionnelle	92
7.4	Modèle radiométrique spatio-temporel	94
7.5	Minimisation de la longueur de description totale	94
7.5.1	Présentation de l'algorithme	95
7.5.2	Estimation des multi-fusions	96
7.5.3	Multi-Fusion et mise à jour de la pile de priorité	100
7.6	Conclusion	101

8	Analyse géométrique du graphe d'adjacence temporelle	103
8.1	Importance du contexte	104
8.2	Algorithme	104
8.2.1	Correspondances 1-1	105
8.2.2	Correspondances 1-n	106
8.2.3	Post-traitement	106
9	Recherche de séquences de caractéristiques fréquentes dans le graphe d'adjacence spatio-temporelle	109
9.1	Classification des caractéristiques TL des régions	111
9.1.1	Fréquence d'un motif dans le LDSOTAG	114
9.1.2	Les méthodes de recherche de motifs fréquents	114
9.1.3	ECLAT pour les graphes	115
9.1.4	Ordre des motifs	119
9.2	Interprétation théorique des évolutions fréquentes	120
9.2.1	Superposition	120
9.2.2	Effet des variations d'évolutions intra-classes, et quantification vectorielle	121
9.2.3	Effet du croisement d'évolutions	121
9.2.4	Effet d'erreurs de segmentation ou d'appariement	121
9.2.5	Regroupement par prototypes	122
9.3	Réglage de la longueur et de la fréquence des motifs pertinents	124
9.3.1	Conclusion	124
V	Expérimentations et résultats	127
10	Validation	129
10.1	Segmentation spatiale sous contrainte temporelle	129
10.1.1	Prétraitements	129
10.1.2	Résultats	130
10.2	Construction du SOTAG	134
10.3	Recherche de motifs fréquents dans le SOTAG	134
10.3.1	Génération des séquences simulées	140
10.3.2	Résultats pour les séquences simulées	142
VI	Conclusion et perspectives	149
10.4	Conclusion	151
10.5	Perspectives	153
A	Annexes	155
A.1	Le principe de minimisation de longueur de message (MDL)	155
A.2	Etat de l'art de l'analyse de séquences d'images	156
A.2.1	Séquences d'images de fluides	156
A.2.2	Séquences d'images médicales	157
A.2.3	Séquences vidéo multimedia	158
A.2.4	Mouvement dans les images satellitaires de la terre	162
A.2.5	Mise en correspondance d'images	172

A.3 Recherche de motifs fréquents	174
A.3.1 Apriori	175
A.4 Algorithmes en profondeur	176
A.5 Croissance-FP	176
A.5.1 Représentations condensées	176
A.5.2 Recherche d'ensembles de symboles maximaux	178
A.5.3 Recherche de motifs séquentiels fréquents	178
A.6 Recherche dans des graphes	179
A.6.1 Recherche de traversées fréquentes dans des graphes	180
A.7 Généralités sur les graphes	182
Bibliographie	185
Publications	209
Glossary	211
Acronyms	213
Notations	215

Introduction :

0.1 Intérêt des STIS

L'ajout de la dimension temporelle à l'observation satellitaire de la Terre ouvre un grand nombre d'applications. Par exemple, pour l'analyse d'écosystèmes. En effet, l'activité humaine a profondément affecté les écosystèmes, et la détection et l'analyse des changements suscite un intérêt grandissant. Ainsi, les STIS sont selon [Pettorelli et al., 2005] une *mine d'or* en ce qu'elles sont des données à grande échelle temporelle et spatiale contrairement aux données terrain habituellement utilisées dans ce cadre. Les STIS peuvent aussi être utilisées pour évaluer l'influence d'une guerre, de changements politiques, d'une sécheresse, d'un feu, d'une inondation, ou encore d'un gel sur l'occupation du sol, les pratiques agricoles, etc [de Beurs and Henebry, 2004, 2008, Kawamura et al., 2004, Wang et al., 2003b, Tait and Zheng, 2003]. Au delà de l'analyse des changements, un suivi temporel d'une zone permet la prédiction de phénomènes qui peut être utile par exemple pour déclencher des alarmes ou identifier des zones à risque [Ceccato, 2004], ou encore dans un cadre écologique afin de prédire la biodiversité [John et al., 2008]. Le biome, le taux de CO_2 atmosphérique et le climat sont très liés et s'influencent mutuellement. La surveillance des forêts au cours du temps présente donc un intérêt particulier [Piwowar et al., 2006, Bartalev et al., 2000, Ito et al., 2007]. La surveillance de l'évolution de cultures fournit une assistance importante aux agriculteurs de grands domaines. Par ailleurs, l'observation de démolitions, de constructions ainsi que de leur évolution est une information importante pour les urbanistes et les autorités régionales. En particulier pour les villes de forte croissance pour lesquelles les constructions précèdent parfois les statuts d'occupation du sol prévus par les autorités. De plus, une prédiction de ces évolutions serait utile pour des organismes publics ou des entreprises de mise en place d'infrastructures nécessaires aux agglomérations d'une certaine taille comme le réseau routier, ferré, téléphonique, internet, d'eau, etc. Enfin, ce type de séquence peut s'avérer d'un intérêt capital dans le cas de crises. En effet, l'observation régulière de la Terre permet un apprentissage des évolutions et changements *normaux* et permet donc par complémentarité de détecter les changements anormaux tels qu'un feu, une inondation, un tsunami, un tremblement de Terre, une migration de population, etc.

Les satellites des générations récentes ont des propriétés optiques permettant d'acquérir des images à très haute résolution. De plus, leurs propriétés mécaniques leur permettant de suivre une scène lors de leur rotation autour de la Terre, ainsi que leur utilisation en constellation, permet une revisite fréquente d'un même site. Un nouveau type de données est donc maintenant accessible : les séquences temporelle d'images satellitaires à haute résolution (STIS-HR). Ces données sont riches en information, mais en contrepartie très complexes ce qui rend leur interprétation visuelle laborieuse. Des analyses automatiques de ce type de données sont donc requises.

0.2 Position du problème

L'analyse automatique s'appuie toujours sur un modèle de données. Nous avons dû dans un premier temps trouver un modèle pour les STIS. Nous avons, dans ce but étudié la littérature sur la modélisation de données 3D et 2D+t. Dans cet état de l'art, détaillé dans la partie I, aucun des modèles proposés ne paraît adapté aux STIS. En effet, les STIS-HR sont des séquences d'images bien particulières. La haute résolution spatiale qui les caractérise (20m pour les images de la STIS ADAM décrite en détail dans la partie II sur laquelle nos algorithmes seront testés, 5m pour les STIS-HR acquises par la constellation *RapidEye*, et 70cm pour les images qu'acquerra la constellation *Pléïades*) permet l'identification d'objets

sémantiques. C'est ce point qui rend les STIS-HR particulières par rapport à des STIS à basses et moyennes résolutions. Une description de la scène dynamique par les objets qui la composent et leur évolution au cours du temps paraît alors naturelle.

A chaque instant, ces objets sont caractérisés, comme pour une scène statique, par une zone spatiale radiométriquement cohérente. Au cours du temps, leur forme reste généralement inchangée. Cependant, leurs radiométries évoluent indépendamment les unes des autres provoquant des changements topologiques. Ainsi, au printemps, lors de la levée des cultures, on observe une division des zones agricoles en différentes parcelles, et symétriquement, à l'automne, une fusion des parcelles lors des récoltes. Ces évolutions radiométriques provoquant des changements topologiques sont caractéristiques des STIS-HR, et il nous faut donc les prendre en compte dans le modèle pour une description fidèle aux données. Par ailleurs, le parcellaire change au gré des exploitants, et un suivi automatique de ces changements présente un intérêt en soi. On appelle scène la zone observée. La scène est délimitée par une étendue spatiale, et une durée. La scène est à distinguer de la STIS qui fait référence à la séquence d'images. Nous proposons, au chapitre 4, d'abord un modèle de la scène permettant de prendre en compte ces changements particuliers, puis nous modélisons les STIS dont le modèle est différent en raison des conditions d'acquisitions qui rendent une partie de l'information inaccessible. On considère que la scène est constituée d'objet spatio-temporel (STO, Spatio-Temporal Object) qui évoluent au cours du temps et interagissent. La modélisation de la scène nous permet de formaliser ce que l'on souhaiterait idéalement déduire des acquisitions. Cependant, les conditions de l'acquisition ne permettent pas d'accéder à toutes les informations nécessaires à la reconstruction d'un modèle aussi complet. En effet, les données sont généralement temporellement sous-échantillonnées par rapport aux phénomènes observés. Par exemple, un labour est un phénomène continu sur une durée de moins d'une journée alors qu'on observe la scène au mieux tous les jours. De plus, les changements de conditions atmosphériques entre les différentes acquisitions provoquent des erreurs d'inter-calibration. Ces deux contraintes nous empêchent de reconstruire les modèles continus sous-jacents. Enfin, des objets (nuages ou avion par exemple) peuvent occulter temporairement des parties d'objets de la scène. Nous restreindrons donc l'étude à l'arrière-plan de cette scène complexe qui comprend des objets immobiles mais dont les caractéristiques radiométriques évoluent dans le temps entraînant des changements topologiques. La structure proposée pour modéliser l'arrière-plan est un graphe d'adjacence temporelle d'objets spatiaux (SOTAG, Spatial Object Temporal Adjacency Graph) dont les nœuds représentent les objet spatiaux (SO, Spatial Object). Ces SO sont liés temporellement, et un parcours des chemins de ce graphe permet de suivre les évolutions des objets au cours du temps. De plus, en considérant une version attribuée de ce graphe, on peut classer les évolutions des STO qui correspondent dans notre représentation aux chemins du graphe. Classifier ces évolutions permet d'identifier différentes natures d'objets. Les STO sont en effet caractérisés par leur évolution radiométrique, ainsi, une culture de colza peut être caractérisée par la séquence de radiométries associées successivement à la terre nue, aux feuilles, aux fleurs, et aux reliquats de culture post-récolte. La figure 1 illustre les différentes étapes de la méthode proposée.

La scène ainsi modélisée, il nous faut ensuite extraire le graphe des données c'est-à-dire déterminer les SO dans chaque image, et mettre ces derniers en relation. Déterminer les SO correspond à une étape de segmentation des images. Il existe de nombreux algorithmes de segmentation, cependant, appliquer le même algorithme de façon indépendante à chacune des images ne permet pas de mettre en évidence les mêmes objets aux différents instants, et il en résulte un SOTAG très complexe. En effet, les conditions d'acquisition changent entre

différentes prise de vues, et les changements simultanés et inconnus de la scène compliquent l’(inter)-calibration des images. La dynamique des différentes images varie donc et de plus, les contrastes locaux entre régions voisines évoluent en raison des évolutions radiométriques propres des différents objets. Les méthodes classiques de segmentation spatio-temporelle ou volumique ne sont pas non plus adaptées aux STIS en raison des variations radiométriques éventuellement importantes entre les différentes acquisitions d’un STO au cours du temps. Nous proposons donc un algorithme de segmentation adapté aux STIS permettant de mettre en évidence des objets spatiaux correspondant à des descriptions comparables des différentes images. Hormis des divisions et fusions occasionnelles, la forme des STO reste généralement inchangée. L’estimation du graphe nécessite donc une estimation conjointe des arcs et des nœuds du graphe car l’estimation des correspondances entre les SO suppose que ces derniers sont connus, et la validité des SO s’appuie sur une bonne correspondance. La description la plus adaptée à la scène est donc la description associée au graphe le plus simple tout en restant fidèle aux données. Ce compromis entre fidélité et simplicité du modèle est typique des problèmes de modélisation, et de nombreuses approches permettant de trouver le compromis optimal existent. En particulier, le principe de minimisation de la longueur de description (MDL, Minimum Description Length) permet de trouver un compromis optimal et sans paramètre. Le problème est alors vu comme un problème de transmission : afin de transmettre la scène, on transmet un message en deux parties dont la première est la description des données en utilisant son modèle, et la deuxième la description du modèle. Dans notre cas, le modèle à décrire est le graphe. Sa description consiste en la description de ses nœuds, c’est-à-dire de la forme des SO associés, ainsi que leur radiométrie. La forme des SO est décrite de façon conditionnelle en s’appuyant sur les correspondances entre les nœuds du graphe. La radiométrie, en revanche, est décrite de façon indépendante pour chaque SO de façon à considérer un *a priori* non informationnel sur les évolutions radiométriques permettant des discontinuités radiométriques. L’estimation conjointe des nœuds et des arcs est néanmoins très complexe, et nous proposons de simplifier le problème dans un premier temps en ne considérant que des correspondances 1-1 pour les descriptions conditionnelles des formes. Cette approximation favorise une description présentant peu de correspondances 1-n, n-1, et surtout n-m et est donc bien adaptée aux STIS. Cette étape, que nous avons appelée *segmentation jointe* est décrite en détail au chapitre 7. Cette segmentation jointe permet d’obtenir une séquence de partitions comparables à partir desquelles on peut par la suite déduire l’ensemble des correspondances. A l’issue de cette partie, décrite au chapitre 8, nous disposons du modèle de STIS, et c’est après ces deux étapes seulement que la STIS peut être analysée par ce modèle.

L’analyse que nous proposons ici est une analyse radiométrique. Les chemins du SO-TAG permettent un suivi des objets dans le temps, et décrivent donc la dynamique des STO. En associant aux nœuds les paramètres du modèle radiométrique de la région spatiale associée, on obtient pour chaque chemin, une séquence de paramètres radiométriques correspondant à l’évolution radiométrique d’un STO. Afin de proposer un traitement automatique du début à la fin, nous avons considéré une analyse automatique de ces séquences. Il existe des méthodes d’analyse automatique de signaux temporels. Cependant, acquérir régulièrement des images à ces résolutions spatiales est un réel défi technologique. On arrive néanmoins à acquérir des STIS à fréquence temporelle d’acquisition mensuelle pour ADAM et journalière pour RapidEye et Pléiades. Cependant les fréquences caractéristiques des objets visibles à ces résolutions spatiales peuvent être très élevées, et les conditions de Shannon sont donc rarement respectées pour les phénomènes d’intérêt. Pour cette raison, nous avons écarté les types de méthodes supposant une *certaine* continuité temporelle des

signaux. De plus, la scène qui nous intéresse est la scène au sol, or cette scène est superposée à l'atmosphère qui présente elle aussi une dynamique spatio-temporelle. Cette évolution radiométrique indépendante de la scène complique encore l'estimation de l'évolution radiométrique des objets d'intérêt. Nous proposons donc de nous placer dans un espace discret en considérant les différentes radiométries discrétisées associées à chaque SO comme des symboles sans relation d'ordre. De cette façon, on limite la comparaison entre deux attributs radiométriques à une comparaison binaire : égalité ou différence. L'algorithme proposé s'inspire de la recherche de motifs fréquents dans des bases de données séquentielles dont un état de l'art est fourni au chapitre 6. L'algorithme s'appuie sur un parcours en profondeur de l'arbre de préfixe décrivant l'ensemble des motifs fréquents et une énumération efficace des motifs dans le graphe par de simples opérations logiques entre octets représentant les séquences associées aux chemins du graphe. Une méthode d'énumération adaptée à la structure du graphe considéré s'appuyant sur une distinction entre deux types de sous-chemins est proposée. De cette façon, on évite une sur-estimation du nombre d'occurrences de certains (petits) motifs liée au fait qu'ils correspondent à des portions de graphes communes à plusieurs chemins.

Ce manuscrit se décompose en six parties : Dans la partie III, un état de l'art général sur la modélisation de données spatio-temporelles et des méthodes d'analyses associées est présenté. Ensuite, les données et le problème sont présentés. Puis nous introduisons le cadre théorique dans lequel s'inscrivent nos travaux. Cette partie se décompose en trois chapitres. Dans le premier, nous introduisons le modèle de STIS que nous proposons et qui sera utilisé par la suite. Puis viennent deux chapitres d'état de l'art, l'un sur la segmentation, et l'autre sur la recherche d'ensembles fréquents permettant d'introduire des concepts, notations, et outils théoriques dont nous nous serviront par la suite, respectivement dans les chapitres 7 et 9. Le cadre théorique une fois posé, nous présentons, dans la partie IV, les algorithmes proposés. Cette partie est constituée de trois chapitres consacrés à la segmentation jointe, à la construction du SOTAG, et à l'analyse radiométrique. Enfin, vient une partie d'expérimentation au cours de laquelle les résultats de la validation des algorithmes sur des données simulées et réelles sont présentés. Finalement, dans une dernière partie, nous tirons des conclusions et proposons des perspectives pour ces travaux.

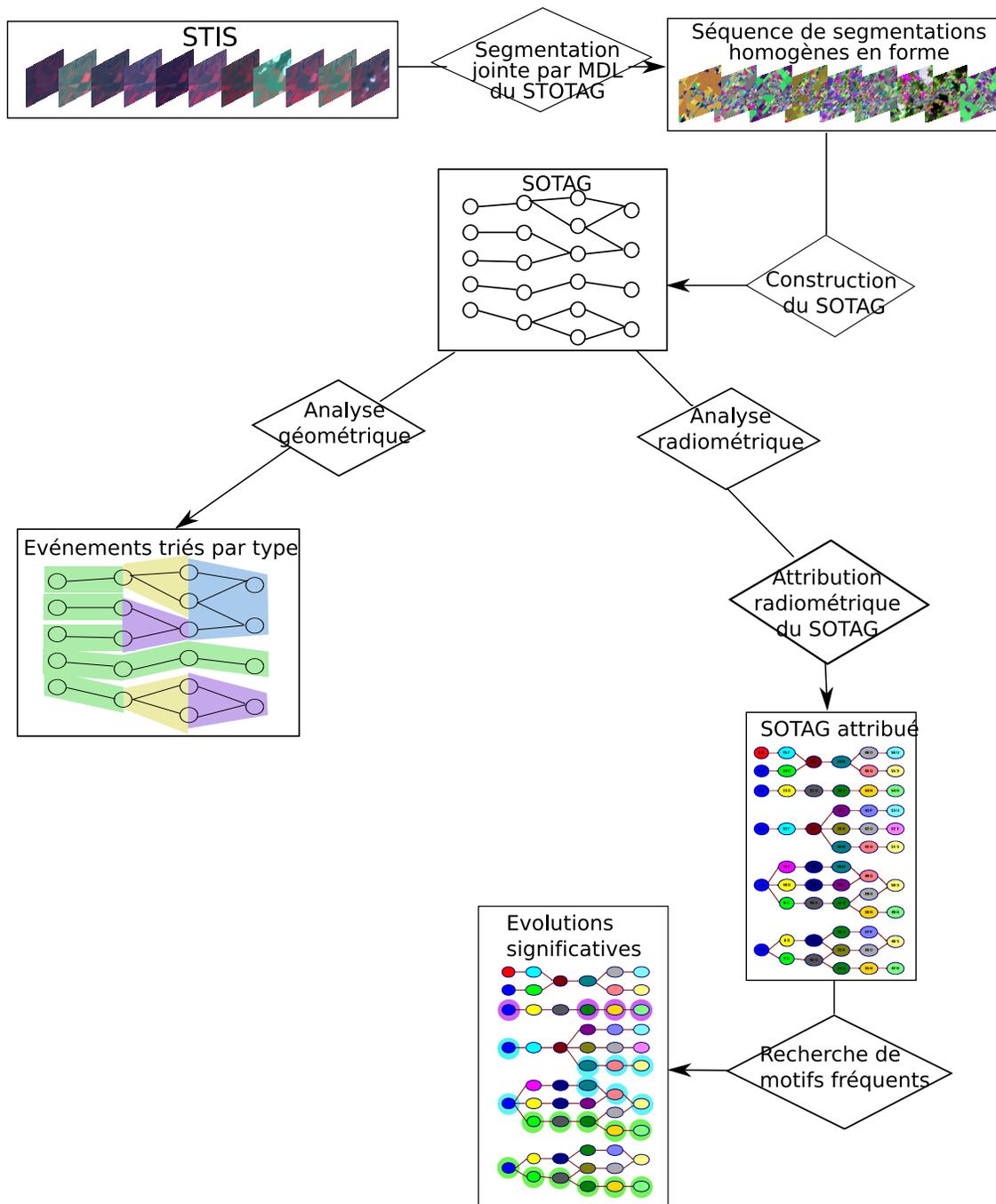


FIGURE 1 – Le traitement que nous proposons est séquentiel. La première étape consiste en une segmentation jointe de images de façon à extraire des objets spatiaux temporellement cohérents. A partir de cette séquence de segmentations, on construit un graphe d’adjacence temporelle des objets spatiaux, le SOTAG. Nous proposons ensuite deux applications de ce graphe : l’une géométrique, classifie les nœuds en fonction de leur degré entrant et sortant traduisant le type de changement géométrique qu’a subit la région. Alternativement, une analyse des évolutions radiométriques est proposée. Pour cela, les nœuds du SOTAG sont d’abord attribués selon les valeurs radiométriques des régions correspondantes, puis, des motifs fréquents sur les chemins du graphes sont extraits.

Première partie

Etat de l'art

Chapitre 1

Différentes séquences, différentes méthodes d'analyse et différents modèles

Une séquence d'images est un ensemble d'images acquises de façon à étudier un espace de représentation plus grand. Ainsi, une séquence d'images représentant des coupes à différentes hauteurs permettent d'étudier un espace 3D. Dans le cas de séquences temporelles, des images d'une scène 3D sont acquises à différents instants de façon à constituer des données $2D+t$.

L'objectif de cette thèse est d'analyser une STISHR. Pour ce faire, il nous faut dans un premier temps modéliser ces données, de façon à pouvoir ultérieurement estimer ce modèle. Il n'existe pas de modèle pour ces données particulières. Cependant, il existe de nombreux modèles adaptées à différentes séquences d'images. Dans ce chapitre, nous passons en revue ces différents modèles de façon à mettre en évidence les ressemblances et les différences avec nos données. Plus de détails, ainsi que les références sur cet état de l'art pourront être trouvés en annexe ??.

1.1 Position de notre problème par rapport à l'état de l'art

En fonction du domaine, le type de changements des objets diffère et des techniques adaptées s'appuyant sur des hypothèses d'homogénéité sont utilisées. Ainsi, les méthodes d'analyse de vidéo supposent généralement un mouvement rigide éventuellement articulé dont l'estimation s'appuie généralement sur une hypothèse d'invariance radiométrique. Quelques approches autorisant de plus une déformation des objets, ainsi qu'une variation radiométriques ont néanmoins été proposées. Ces méthodes s'appuient sur une mise en correspondance de points caractéristiques. Ces points caractéristiques supposent une invariance locale du contraste, or une telle propriété n'est plus vérifiée pour les STIS.

D'autres types de séquences présentent des évolutions radiométriques locales. C'est le cas par exemple des séquences de fluide, des séquences d'images médicales, et des séquence temporelle d'images satellitaires à résolution grossière, *coarse resolution* (STIS-CR). Pour les fluides, une approche pixellaire et un *a priori* provenant d'une connaissance physique du phénomène permettent d'estimer les mouvements des particules malgré les variations radiométriques. Cette approche n'est doublement pas applicable aux STIS-HR. D'une part, en raison de notre ignorance des phénomènes observés liée à l'hétérogénéité de la scène,

et par le fait que cette approche ne permet pas de caractériser des objets en raison de sa nature pixellaire.

Les séquences d'images médicales peuvent présenter des déformations. En conséquence, une hypothèse d'invariance radiométrique est dans un premier temps formulée de façon à estimer le mouvement rigide puis non rigide des objets. Puis, pour l'analyse radiométrique des séquences médicales (comme des STIS-CR) des signaux temporels pixellaires peuvent être considérés. Ces signaux peuvent être filtrés, en supposant un modèle, par filtrage local ou dans l'espace fréquentiel. Ces approches s'appuient sur des hypothèses fortes de continuité ou d'*a priori* que nous ne pouvons formuler pour les STIS en raison du sous-échantillonnage temporel par rapport aux évolutions et de l'hétérogénéité de la scène.

Des caractéristiques peuvent ensuite être extraites de ces signaux de façon à les regrouper par similarité. Dans ce cas, cependant, aucune contrainte de connexité spatiale n'est supposée, et les objets sont donc implicites, et nécessitent un post-traitement pour être identifiés. L'approche inverse s'appuyant sur une partition spatiale dont chaque composante est caractérisée par un signal temporel que l'on souhaite analyser existe aussi. Enfin, des approches de détection de changements s'appuyant éventuellement sur une segmentation ont aussi été proposées. Toutes ces approches supposent cependant une partition spatiale unique, et un support temporel des évolutions fixé et compact. Ce sont ces deux points que nous proposons de remettre en question lors de l'analyse de STIS-HR en supposant des changements topologiques de type fusion et division ainsi que des évolutions radiométriques ayant une durée de vie et pouvant être perturbées par des occultations.

Enfin, dans le domaine de la vidéo, les méthodes d'analyse classiquement employées permettent d'extraire deux types de modèles de scènes spatio-temporelles : des scènes constituées d'un ensemble de séquences temporelles d'objets spatiaux et des scènes constituées de séquences d'un ensemble d'objets spatiaux. Les objets spatiaux considérés peuvent présenter une homogénéité radiométrique, de mouvement, de profondeur, sémantique et les séquences d'objets spatiaux peuvent être caractérisées par une redondance temporelle de mouvement, de radiométrie, de forme ou encore une combinaison de ces caractéristiques.

Deux types de mouvements sont considérés pour les objets : les mouvements rigides et les mouvements de déformation. L'estimation de mouvement est généralement basée sur une homogénéité temporelle de mouvement par une optimisation itérative. Mais elle peut aussi s'appuyer sur une homogénéité radiométrique, de forme, de points caractéristiques, ou encore combiner ces différentes approches. De petites déformations peuvent être prises en compte de façon implicite en considérant des homogénéités radiométriques temporellement localisées, et des déformations plus importantes peuvent être prises en compte avec des méthodes basées modèle.

Certaines approches de poursuite d'objet ou de détection de changement prennent en compte les événements de type apparition. En revanche, à notre connaissance, les événements de type fusion et division, qui sont des phénomènes observés dans nos données, ne sont pas pris en compte dans les modèles proposés dans ce domaine,¹.

La scène que nous considérons est beaucoup plus générale, et présente une variété de changements plus importante par rapport aux modèles de scène sous-jacents des méthodes sus-mentionnées. Nous proposons donc un modèle de STIS permettant de prendre

1. En annexe, nous présentons différents cas de traitement vidéo où des fusions et des divisions d'objets devaient être gérés, cependant, il s'agissait dans ces cas de conflits liés à la modélisation 2D d'une scène 3D. Ces conflits étaient d'ailleurs généralement gérés en considérant un ordre de profondeur pour les objets. Les changements topologiques que nous considérons pour les STIS ne sont pas liés aux occultations, mais aux changements radiométriques des différents objets.

en compte cette hétérogénéité des objets spatio-temporels (OST). Dans ce cadre, nous proposons une brève revue des modèles de bases de données spatio-temporelles, et particulièrement des systèmes d'information géographiques temporels pour lesquels les problèmes de changements topologiques et d'identification qui en découlent ont été abordés.

Le tableau 1.1 synthétise l'état de l'art détaillé en annexe en mettant l'accent sur le type de modèle appliqué aux données.

	Homogénéité spatiale					Homogénéité temporelle					OS	OST	Fus/Div		
	Rad	Mouv		F	P	Sem	Rad	Mouv		F				P	Sem
		R	D					R	D						
[Diehl, 1991]			✓		✓		✓		✓				✓		
[Gelgon and Bouthemy, 2000] [Wang, 1998]			✓				✓		✓				✓		
[Wang and Adelson, 1994]			✓				✓		✓				✓		
[De Roover et al., 2005]	composite	~	✓						✓				✓		
[Marcotegui, 1996]	composite	✓					✓							✓	
	✓						✓							✓	
	composite	✓							✓		✓			✓	
[Meier and Ngan, 1998]	objets composites, mise en correspondance de points caractéristiques												✓		
[Moscheni et al., 1998]	✓								✓					✓	
[Bremond and Marquès, 1996]	✓						Evolution de paramètres de forme						✓		
[Gu and Lee, 1997]	composite					✓	✓	✓	✓					✓	
[Meyer and Bouthemy, 1993, 1994]			✓				✓		✓					✓	
[Bouthemy and Francois, 1993]	✓						✓							✓	
[Eisert and Girod, 1998]	✓						✓	✓		✓				✓	
Modèle pour la STIS	✓									✓	✓			✓	✓

TABLE 1.1 – Comparaison des modèles d’objets existant dans l’état de l’art et de la modélisation des objets de la scène. Le symbole ✓ signifie que l’homogénéité est directement exploitée pour les objets, et le symbole √ signifie qu’elle est exploitée indirectement (pour l’estimation du mouvement par exemple). Rad=radiométrique, Mouv=mouvement, R=déformation, D=déplacement, F=forme, p=profondeur, Sem=sémantique, Fus=fusion, Div=division. La mention *composite* signifie que les méthodes autorisent de petites déformations du contour extérieur, en exploitant une homogénéité radiométrique spatiale des composantes des objets.

Chapitre 2

Modélisation de données spatio-temporelles

2.1 Modèle de base de données spatio-temporelles

La prise en compte du temps dans des bases de données spatiales est un sujet d'intérêt depuis une vingtaine d'années. En effet, les données considérées peuvent être évolutives, et il est intéressant dans ce cas de mettre à jour la base de données, ou encore d'analyser les changements. Différents modèles de bases de données spatio-temporelles existent. Ces modèles permettent de stocker et de manipuler des données spatio-temporelles. Il paraît donc intéressant d'effectuer un état de l'art sur ces modèles dans la mesure où l'on souhaite construire un modèle pour des STIS qui sont des données spatio-temporelles.

2.1.1 Modélisation du temps

Différents types de modélisation du temps existent. Le temps peut être linéaire, cyclique ou encore branché [Peuquet, 1994]. Cette dernière modélisation permet de prendre en compte les différentes évolutions possibles dans le passé ou dans le futur.

Par ailleurs, le temps dans une base de données est multidimensionnel, chaque dimension correspondant à un des aspects revêtus par le temps. Ces différents aspects sont le temps de validité, le temps de transaction, et le temps d'affichage [Snodgrass, 1992]. Le temps de validité concerne le temps auquel un fait est vrai dans la réalité par opposition au temps de transaction qui concerne le temps auquel le fait a été stocké dans la base de données. Le temps de transaction d'un événement est un intervalle entre le moment où des transactions liées à cet événement ont inséré de l'information dans la base de données, et le moment où cette information a été retirée. Un exemple d'événement ayant un temps de transaction mais pas de temps valide est une proposition d'achat ou un plan d'aménagement. Ces événements existent effectivement et peuvent donc être enregistrés dans la base de données, cependant, ils n'existent pas dans le monde réel et ne peuvent par conséquent être acquis ou mesurés. Une base de données contenant seulement le temps de validité, est qualifiée d'*historique*, si elle contient seulement le temps de transaction, on l'appelle *rollback* (car elle permet de revenir à une version précédente), et si elle contient à la fois le temps de validité et le temps de transaction, bitemporelle.

Dans le modèle temporel orienté objet présenté dans [Allen, 1983], sept relations temporelles sont proposées :

- Avant : $A < B \Leftrightarrow t_A^e < t_B^b$

- Egal : $A = B \Leftrightarrow t_A^b = t_B^b \& t_A^e = t_B^e$
- Précède : $AmB \Leftrightarrow t_A^e = t_B^b$
- Superpose : $AoB \Leftrightarrow t_A^b < t_B^b < t_A^e \& t_A^e < t_B^e$
- Pendant : $AdB \Leftrightarrow t_A^b < t_B^b < t_A^e \& t_A^e > t_B^e$
- Commence : $AsB \Leftrightarrow t_A^b = t_B^b$
- Finit : $AeB \Leftrightarrow t_A^e = t_B^e$

Événement, état, et processus Un événement est quelque chose qui arrive ; un événement a un début et une fin, et les événements sont dénombrables [Galton, 2001]. Par exemple l'événement "je vais en vélo à l'université". Cet événement arrive plusieurs fois dans l'année et on peut compter le nombre de fois où ça arrive.

Un état ou un processus est quelque chose qui est vrai à tout instant (pendant la durée de validité). Par exemple, "je suis à l'université". Pendant toute la durée où "je suis à l'université", la propriété est vraie. On note ici la différence avec un événement : on ne peut pas dire que pendant toute la durée de l'événement "je vais à vélo à l'université", la proposition a lieu. L'occupation du temps par un état ou un processus est plus homogène.

On peut donc modéliser le temps par événements ou par états et processus (modélisation *fluent*). Cette distinction de modélisation temporelle est similaire à la distinction entre le modèle par champs et le modèle orienté objet dans le cas de données spatiales. Ainsi, de même que la modélisation par champs consiste à attribuer à chaque position un scalaire, la modélisation par fluent consiste à associer à chaque temps un scalaire. Et la modélisation par objet consiste à attribuer à chaque objet une localisation et des attributs, et la modélisation par événements consiste à attribuer à chaque événement un temps et des attributs [Worboys, 1994].

Granularité temporelle On peut considérer trois granularités du temps [Snodgrass, 1992] :

- discret : temps isomorphe à \mathbb{N}
- dense : entre deux instants, il existe un instant. Dans ce cas, le temps est isomorphe à \mathbb{R} ou à \mathbb{Q}
- continu : temps isomorphe à \mathbb{R} (dense sans "trou")

Pour des modèles continus, l'unité temporelle n'a pas de durée. Dans le cas de modèles discrets, en revanche, l'unité temporelle, appelée chronon a une durée qui est la granularité du temps.

Le choix du chronon dépend de la nature des phénomènes considérés [Laurini and Thompson, 1992] Il existe en effet différentes échelles temporelles. Ainsi, une montagne peut paraître immobile de notre point de vue bien que sa forme évolue à l'échelle temporelle géologique en raison de l'érosion.

2.1.2 Modélisation de l'espace spatio-temporel

Le concept de temps implique que des changements apparaissent dans le présent, le passé et le futur de la durée de vie d'un phénomène du monde réel. Un système d'information géographique temporel a pour but de comprendre ces changements et leurs effets au cours du temps plutôt que de les reproduire simplement en affichant une séquence de captures. [Wachowicz and Healey, 1994]

Une base de données spatio-temporelle a plusieurs dimensions : la dimension thématique, la dimension spatiale, et la dimension temporelle.

L'analyse de données spatio-temporelle nécessite d'une part la construction de bases de données spatio-temporelles, et d'autre part un langage de requêtes pour ces bases de données. Plusieurs modèles de base de données ont donc été proposés dans ce but.

[Snodgrass, 1992] propose cinq cas à considérer pour la variation spatio-temporelle d'attributs :

- Evolution de l'attribut d'un objet spatialement invariant. C'est la *temporalité de l'attribut*. La population mondiale en est un exemple. La spatialité est alors superflue : des systèmes de gestion de données temporelles suffisent à traiter ce cas.
- La valeur d'un attribut associé à une région spatiale peut varier spatialement, mais pas temporellement. L'altitude par exemple varie spatialement, mais pas temporellement. La temporalité est alors superflue : un GIS suffit à traiter ce cas.
- La valeur d'un attribut peut évoluer dans le temps et dans l'espace. C'est le cas par exemple du pourcentage spatial de couverture nuageuse.
- Les frontières des objets peuvent évoluer dans le temps. On parle alors de *topologie temporelle*. Les attributs thématiques des objets peuvent néanmoins être invariants dans le temps.
- Un objet qui évolue et dont les attributs thématiques évoluent.

Dans les deux derniers cas, l'intégration de la dimension temporelle aux GIS est indispensable.

[Claramunt and Thériault, 1995] proposent une classification plus détaillée des deux derniers cas sus mentionnés en trois catégories : les changements concernant l'évolution d'une simple entité, des changements concernant les relations fonctionnelles entre entités, et les relations concernant l'évolution de structures spatiales impliquant plusieurs entités. Ces changements sont décrits figure 2.1, et les évolutions du troisième type seront utilisées dans cette thèse.

2.1.3 Modèle de base de données spatio-temporelles

Un modèle de base de données (ou modèle de données) est un ensemble de concepts et de règles permettant de décrire les structures et les caractéristiques d'intérêt par rapport au monde réel considéré. Il permet de spécifier les propriétés structurelles d'une base de données et, de façon idéale, doit fournir un langage permettant à l'utilisateur et au concepteur de la base de données d'exprimer leurs besoins tout en permettant leur implémentation [Worboys et al., 1990]. Ce langage permet de formuler des requêtes. Il doit être générique, complet, correct, extensible, et fournir un ensemble d'opérateurs.

D'après [Abraham and Roddick, 1999], l'utilisateur de données spatio-temporelles veut obtenir des réponses à des requêtes sur l'état de la base de données à différents instants, la localisation d'objets et de phénomènes, leur histoire, l'occurrence d'événements, etc. Ces requêtes sont de trois types : spatiales (localisation, propriétés, et relations), temporelles (temps, propriétés, et relations) et spatio-temporelles (comportement et relations). Les requêtes spatio-temporelles sont classées en trois catégories : les requêtes simples (état d'une entité à un instant), les requêtes sur des durées spatio-temporelles ("qu'arrive-t-il à tel objet sur telle période?"), et les requêtes de comportement (vitesse d'un objet, vitesse maximale, distance parcourue, expansion maximale, etc).

Il existe différents modèles de base de données, allant du plus simple, permettant de répondre à des requêtes concernant exclusivement les données stockées [Langran and Chrisman, 1988] à d'autres, plus complexes, permettant de répondre à des requêtes nécessitant des raisonnements ou des combinaisons des objets stockés.

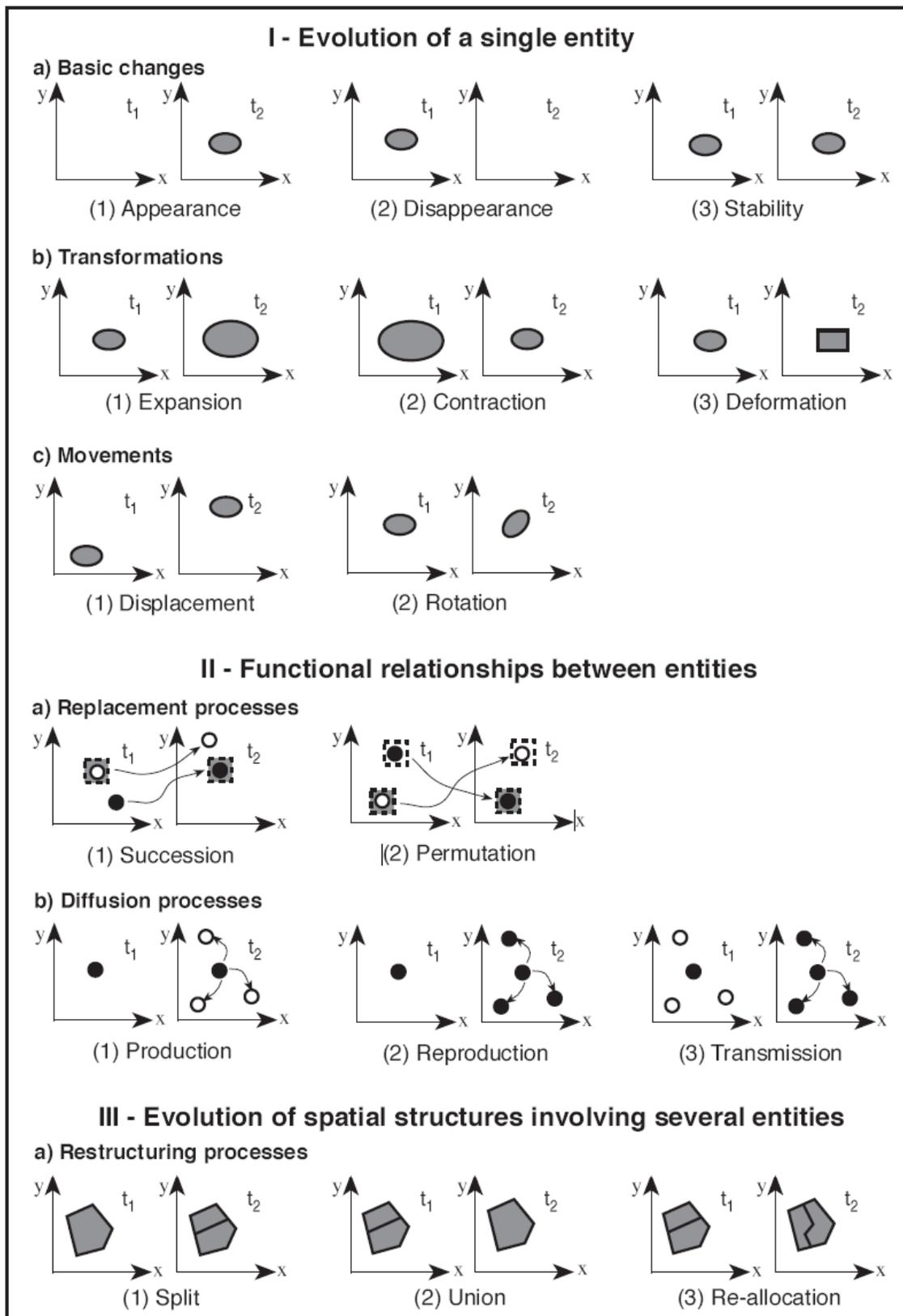


FIGURE 2.1 – Typologie des processus spatio-temporels proposée dans [Claramunt and Thériault, 1995]

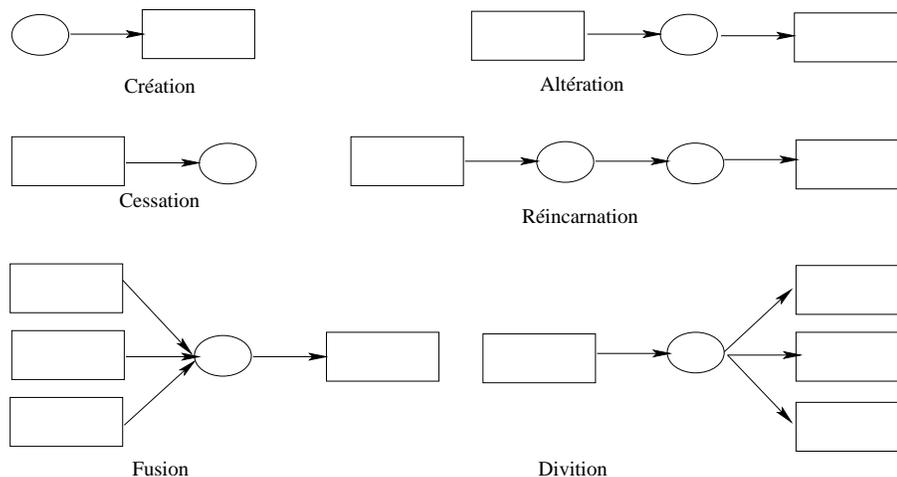


FIGURE 2.2 – Six changements proposés par le modèle de graphe d’histoire de [Renolen, 1996]. Les objets statiques sont représentés par des rectangles de longueur liée à la durée (un état instantané de durée nulle sera représenté par un carré), et les événements sont représentés par des ellipses de longueur liée à la durée (un cercle correspond à un changement soudain). Ces objets constituent les nœuds du graphe d’histoire, et les arcs représentent des relations de succession. Ce modèle considère six types de changements correspondant à différentes contraintes de cardinalité.

Nous avons vu qu’il existe différents types de changements : les événements, et les processus. Certains modèles proposent de modéliser les événements comme des entités à part entière [Peuquet, 1994, Peuquet and Duan, 1995, Frank, 1994, Langran, 1992, Renolen, 1996]. Cette modélisation permet d’une part de répondre facilement à des requêtes concernant des événements, et d’autre part d’accéder à l’histoire de la scène en *rejouant* successivement les différents événements. [Frank, 1994] propose aussi une modélisation par événements, et les événements qu’il considère sont de type apparition et disparition de frontières. Les événements considérés par [Langran, 1992] sont des cartes d’amendements. [Renolen, 1996] propose un modèle, le *graphe d’histoire* permettant de prendre en compte à la fois des événements instantanés, des changements progressifs et des états d’objets. On peut ainsi décrire quatre types de changements : les processus, les états statiques, les événements instantanés, et les événements qui durent. Le graphe d’histoire de la base de données est une représentation graphique de ces changements (voir figure 2.2). Ce modèle permet donc de gérer à la fois des événements et des états. De plus, le temps est décrit de façon à la fois continue et discrète et les changements considérés incluent les déplacements. Enfin, les relations temporelles sont explicites et le modèle répondra donc rapidement à des requêtes temporelles.

L’un des principaux défis pour les bases de données spatio-temporelles est de capturer la persistance d’un objet de façon à pouvoir répondre à des requêtes concernant les évolutions d’un objet. Cependant, au cours du temps, les attributs spatiaux et thématiques d’un objet peuvent évoluer. En conséquence, les identifiants changent, et leur persistance est difficile à capturer.

L’information temporelle a été incorporée aux GIS en ajoutant un attribut temporel aux couches thématiques [Armstrong, 1988], aux attributs [Langran and Chrisman, 1988, Langran, 1992], et aux objets spatiaux [Worboys, 1992]. [Worboys, 1992] considère une partition de l’espace spatio-temporel par des *cylindres* spatio-temporels dont la section peut

changer brutalement. Cette représentation présente une cohérence temporelle contrairement au modèle composite espace-temps (STC) de [Langran and Chrisman, 1988] ou au modèle *snapshot* de [Armstrong, 1988] de sorte que des requêtes concernant les évolutions d'un objet particulier peuvent être traitées. Cependant, comme le note [Yuan, 1999], la cohérence temporelle étant liée à une position spatiale, les requêtes de types mouvement ne peuvent être prises en compte. De plus, les requêtes temporelles sont restreintes aux entités possédant un attribut temporel de sorte que des requêtes simplement temporelles ou simplement spatiales sont plus compliquées.

L'identifiant des objets dépend du point de vue. Ainsi, on peut s'intéresser aux changements thématiques ou aux changements spatiaux en une position spatiale fixée, mais aussi aux changements thématiques ou spatiaux d'une entité. [Yuan, 1999] propose donc de séparer les trois aspects présents dans une base de données spatio-temporelle : le temps, l'espace, et la thématique de façon à pouvoir s'intéresser aux changements de l'un, les autres étant fixés, ou aux changements de plusieurs d'entre-eux simultanément. Le modèle proposé est un graphe constitué de trois types d'objets, spatiaux, sémantiques, et temporels, et d'un arbre représentant les relations d'agrégation des différents objets spatiaux de façon à connaître la persistance des objets spatiaux. Les objets temporels ont des attributs permettant de caractériser des processus, par exemple une transformation, de sorte que les changements brutaux ne soient possibles que par les liens dans le graphe.

[Güting et al., 2000] propose un modèle permettant de prendre en compte les évolutions continues d'objets, et particulièrement les déplacements. Il propose pour cela une généralisation des types de données spatiales à la dimension temporelle. Il définit ainsi le type réel mobile, point mobile, ligne mobile et région mobile. Pour chacun de ces types, il définit des opérations, et un langage de requête associé. On peut ainsi tester les instants pour lesquels un point mobile est à l'intérieur d'une région mobile, ou encore calculer la distance entre deux points mobiles (qui est un réel mobile). Dans ce cas, le temps est modélisé comme une part intégrante des objets.

2.2 Conclusion

Nous avons vu différents modèles de bases de données spatio-temporelles. Ces modèles considèrent différents types de changements, et en particulier, nous avons vu différentes façons de gérer les changements topologiques que subissent les objets dans le temps. Nous retiendrons particulièrement le concept d'événements et de processus permettant de rendre compte de la nature brutale ou continue de changements. Par ailleurs, la plupart des modèles considèrent différentes propriétés essentielles : thématique et spatiale qui peuvent elles aussi évoluer. A notre connaissance, aucun modèle ne prend en compte complètement les changements de type déplacement avec les transformations thématiques, les déformations, les événements topologiques et sémantiques. En effet, dans ce cas, se pose la question de la cohérence des objets à considérer. Cette cohérence dépend du point de vue, aussi, [Yuan, 1999] propose un modèle permettant de retrouver différents types de cohérence temporelles en fonction de la requête. Ces différents points de vue suggèrent donc un modèle dans lequel les identifiants ne sont pas fixés, mais dépendront de l'application. Ainsi, [Raza et al., 1996] introduit le concept de propriété essentielle pour une certaine application déterminant l'existence d'un objet.

Deuxième partie

Présentation des données et du problème

Chapitre 3

Présentation des données

3.1 La STIS ADAM

L'assimilation de données est une technique permettant de réaliser l'estimation des paramètres et des variables d'état d'un système au cours du temps. Cette méthode réalise un lissage de mesures bruitées du système selon un modèle dynamique et une condition initiale.

Dans le cadre du projet Assimilation de Données par Agro Modélisation (ADAM), visant à intégrer les mesures physiques issues des images satellitaires aux modèles physiques utilisés en agronomie, une grande quantité d'images issues de différents capteurs ont été acquises sur une zone d'agriculture intense de la plaine du Danube en Roumanie. Parmi ces données se trouve une STIS à haute résolution. Cette STIS préfigure assez bien les futurs jeux de données que les satellites seront capables d'acquérir. Les algorithmes seront donc validés sur cette STIS.

3.1.1 Les images SPOT

Cette STIS est composée d'images acquises par SPOT 1, 2, et 4 en mode multispectral. Le principe d'imagerie des satellites SPOT est de type barette (push broom). C'est-à-dire que le système observe une ligne au sol, ce qui constitue les lignes d'une image, et les colonnes successives sont obtenues par le déplacement du satellite sur son orbite. Deux systèmes d'acquisition sont associés à ces satellites : les instruments "Haute résolution Visible" (HRV) 1 et 2. Comme leur nom ne l'indique pas, ces systèmes acquièrent non seulement dans le visible mais aussi dans l'infra rouge. Dans le cas du mode panchromatique, un système optique redirige le faisceau vers une barette de 6000 détecteurs. Le champ de vue étant de 4° ce qui correspond à une fauchée de 60km, la résolution résultante est donc de 10m. Dans le cas du mode multispectral, un système optique redirige le faisceau vers un séparateur spectral,¹ qui redirige les 3 (4 pour spot 4) faisceaux résultants vers 3 (4 pour SPOT 4) barettes de 3000 détecteurs. La résolution spatiale résultante est de 20m. Les canaux résultant correspondent au vert (500-590nm), au rouge (610-680nm), et au proche infra-rouge (780-890nm).

Les satellites SPOT sont des satellites qui ont une orbite quasi polaire, circulaire, héliosynchrone à 822 km d'altitude ce qui correspond à un cycle de 26 jours. De plus, ces satellites ont une capacité de dépointage de $\pm 27^\circ$ ce qui leur confère une fréquence de

1. L'utilisation de ce séparateur assure un recalage géométrique des différents canaux

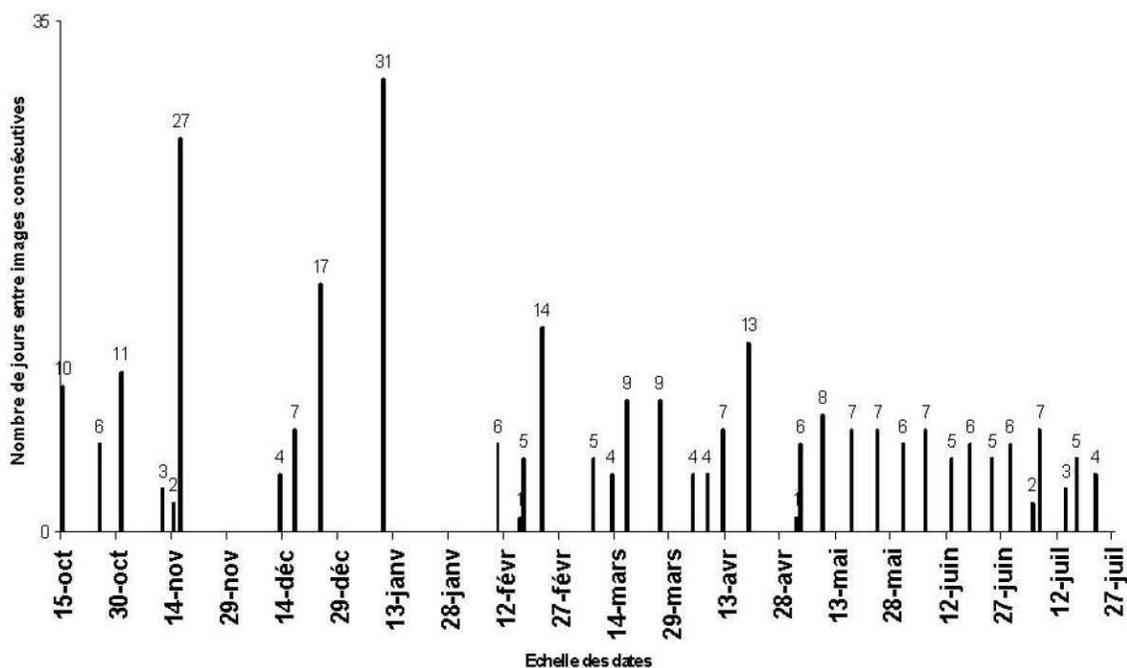


FIGURE 3.1 – Nombre de jours entre l’acquisition d’images sans couverture nuageuse ou neigeuse dans la base de données ADAM. L’axe horizontal représente le temps et supporte l’échantillonnage irrégulier des images. L’axe vertical représente la différence de temps entre les dates consécutives d’acquisition. Notons la baisse du nombre d’acquisition pendant la période hivernale correspondant à des images enneigées ou nuageuses menant à un écart temporel entre images qui atteint 31 jours.

revisite de cinq jours. La combinaison des trois satellites permet d’atteindre une fréquence de revisite quasi journalière.

La durée de l’acquisition de la STIS ADAM s’est étendue sur 286 jours, d’octobre 2000 à juillet 2001. Une étape de sélection des images, de traitements radiométriques et de recalage a ensuite été effectuée pour constituer le jeu de données que nous allons utiliser.

3.1.2 Sélection des images

L’objectif initial de la mission était d’assimiler les données pour de l’agro-modélisation. Des cultures particulières sur un site particulier à l’est de Bucarest étaient observées. Les images pour lesquelles ces zones particulières étaient enneigées ou recouvertes de nuages ont été retirées de la séquence. Il en résulte une STIS de 38 images irrégulièrement échantillonnées dans le temps. La figure 3.1 illustre l’échantillonnage de la séquence.

3.1.3 Prétraitements

Il existe trois niveaux de prétraitements des données SPOT, chaque niveau pouvant être subdivisé en sous-niveaux. La correction de niveau 1A est une correction radiométrique consistant à égaliser la sensibilité des détecteurs. Au niveau 1B, une correction supplémentaire, géométrique est effectuée. Cette correction utilise la trajectoire du satellite de façon à supprimer l’effet panoramique, la rotation et la courbure de la terre, et la variation d’altitude du satellite par rapport à l’ellipsoïde de référence. Au niveau 2A, une correction

Nombre d'images	38
Canal 1	500-590nm
Canal 2	610-680nm
Canal 3	790-890nm
Intervalle temporel maximal entre deux images	31 jours
Intervalle temporel minimal entre deux images	1 jour
Intervalle temporel moyen entre deux images	7.8 jours
Nombre de colonnes	3000
Nombre de lignes	2000
résolution spatiale	20m
Précision du recalage géométrique	1.5 pix

TABLE 3.1 – Spécificités de la STIS ADAM

géométrique supplémentaire est effectuée par projection cartographique de façon à pouvoir combiner l'image avec d'autres informations géographiques. Au niveau 2B, une projection cartographique avec points d'appuis (ou levée GPS sur le terrain) est utilisée, et au niveau 3, la projection cartographique considérée utilise non seulement des points d'appui, mais aussi un modèle numérique du terrain de façon à éviter les distorsions liées au relief à l'aide d'une spatio-triangulation et d'interpolation. Les images de la STIS ADAM sont corrigées au niveau 3 et la précision de ce recalage est de 1.5 pixels.

3.1.4 Correction radiométrique

Le capteur HRV fournit des images de comptes numériques. Ces images sont d'abord transformées en images de luminance, ² à l'aide d'une relation linéaire faisant intervenir le gain du capteur. Les images ainsi transformées sont donc indépendantes du capteur. Cependant, elles restent dépendantes de l'éclairement du soleil qui dépend de la date et de l'heure de la prise de vue, ³. Ces données sont donc prises en compte en calculant le rapport entre la puissance solaire correspondante et l'image de luminance pour les trois plages de longueur d'onde considérées. Les images résultantes sont des images de réflectance "au-dessus de l'atmosphère". Notons qu'en réalité, les surfaces observées ne sont pas lambertiennes, et donc que l'énergie mesurée dépend de l'angle de prise de vue. Considérer des images de réflectances alors que les angles de prises de vue varient constituent donc une première approximation de la correction radiométrique.

Une autre variable importante entre les différentes images est l'atmosphère. D'après [Hagolle et al., 2008], la réflectance au dessus de l'atmosphère au-dessus de cibles sombres peut changer de plus de 100% entre un jour brumeux et un jour clair. Des mesures de température d'humidité, et d'ozone permettent de calculer les réflectances au dessus de la canopée (TC) à l'aide d'un modèle de transfert radiatif. Cependant, en raison du manque de mesures des propriétés optiques des aérosols au moment de chaque prise de vue, de nombreuses approximations ont été effectuées dans cette étape ce qui implique une mauvaise calibration radiométrique des images TC.

Le tableau 3.1 résume les spécificités des données.

2. La luminance est une puissance par unité d'angle solide de visée et de longueur d'onde.

3. L'heure de la prise de vue peut varier bien que le satellite soit héliosynchrone en raison du dépointage

3.1.5 La scène observée : phénoménologie

La STIS ADAM présente différents phénomènes particuliers, et nous proposons dans cette sous section d'interpréter quelques exemples de façon à familiariser le lecteur avec cette STIS.

Décrivons succinctement, ⁴ pour cela dans un premier temps les mécanismes qui influencent les propriétés spectrales de différents objets observables dans les trois canaux utilisés à savoir proche infra rouge, rouge, et vert.

Les plantes utilisent l'énergie solaire pour fabriquer des carbohydrates riches en énergie. Cette énergie solaire est absorbée par des pigments (chlorophylle) dans le rouge et dans le bleu et réfléchi dans le vert, ce qui est responsable de la couleur verte des plantes.

De plus, les feuilles réfléchissent dans le proche infra-rouge. On pourra donc caractériser de la végétation à l'aide de ces deux canaux. Le spectre caractéristique la végétation est représenté figure 3.2 à gauche.

La végétation évolue de façon saisonnière. La description de ces changements saisonniers s'appelle la phénologie. La réflectance typique d'une plante au cours de son développement est décrite par la courbe verte [Belward and Valenzuela, 1991] du graphique à droite dans la figure 3.2. Lorsque les plantes émergent, la réflectance du rouge diminue en raison de l'absorption par la chlorophylle, et la réflectance du proche infra-rouge augmente un peu car la végétation réfléchit dans le proche infra-rouge. Ceci correspond au point (2). Plus la couverture végétale augmente, moins on voit de sol, et donc plus la réflectance du proche infra-rouge augmente. La réflectance du rouge continue de diminuer aussi jusqu'à atteindre un minimum lié au fait que les feuilles sous la canopée sont occultées et donc réfléchissent moins. Ceci correspond au point (3). La réflectance de l'infra-rouge, en revanche continue de croître avec le nombre de feuilles jusqu'au point (4). Lors de la sénescence, la chlorophylle se décompose se qui se traduit par un accroissement de la réflectance dans le rouge, et un changement de l'orientation des cellules des feuilles a pour conséquence une baisse de la réflectance dans l'infra-rouge (points (5) et (6)). Finalement, la culture est récoltée, laissant le sol et les résidus post-récolte (point (7)).

Le comportement que nous venons de décrire correspond à un comportement moyen. Les valeurs des différents points ainsi que les intervalles temporels sont spécifiques à chaque plante et dépendent aussi de la zone géographique et du mode de culture.

Forts de ces considérations, nous proposons différentes interprétations de phénomènes dans la STIS ADAM.

Dans les figures 3.3 à 3.12 différents phénomènes liés à la scène sont présentés. En exploitant la connaissance a priori sur les types de cultures pratiquées et les changements topologiques et radiométriques des parcelles, on peut émettre des hypothèses sur la nature de la culture. La figure 3.3 présente différents exemples de récolte. Les zones rouges correspondent à de la végétation, et les zones sombres à du sol nu. La nature brutale du passage de zone de végétation à zone de sol nu indique une intervention humaine. La figure 3.4 présente différents cas de labours. Ici encore des a priori sur les pratiques agricoles sont utiles pour distinguer une récolte d'un labour sur une terre couverte d'herbes folles. La figure 3.5 montre l'évolution lente de la radiométrie d'une forêt de feuillus, la figure 3.6 un cas de floraison, la figure 3.7 montre une manifestation particulière d'un enneigement, la figure 3.8 présente un cas de brûlis. Et la figure 3.10 illustre l'intérêt d'identifier les labours, récolte, floraison, arosage etc pour l'identification de cultures.

Les figures 3.3 à 3.12 présentent des phénomènes liés à la scène. A ces phénomènes se

4. Pour plus de détails, nous renvoyons le lecteur à [Belward and Valenzuela, 1991]

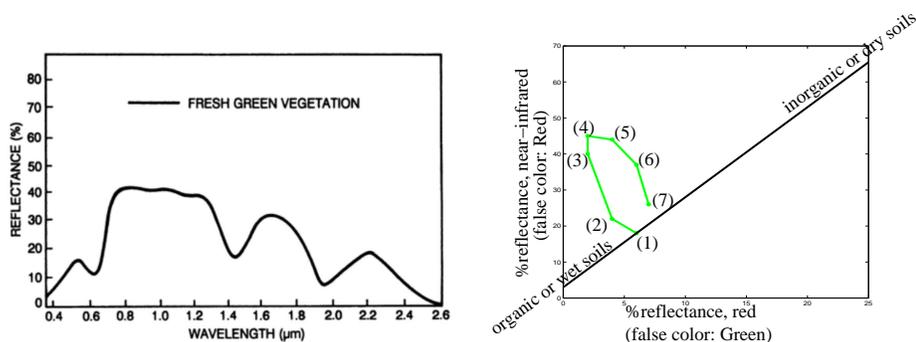


FIGURE 3.2 – a) Spectre de la végétation. Les bandes bleues correspondent aux bandes spectrales des images SPOT constituant la séquence ADAM. b) La courbe verte représente la réflectance typique d'une plante au cours de son développement, la courbe noire les réflectances associées à différents types de sols. Le premier point correspond au sol nu. Ce point peut varier selon la droite noire qui y passe. Près de l'origine, la réflectance correspond à un sol organique ou humide, et loin de l'origine, la réflectance correspond à des sols peu organiques (calcaires par exemple) ou secs.

superposent ceux liés à l'acquisition. Ainsi, figure 3.13, on observe l'effet de la variation de l'angle de la prise qui a pour conséquence que certaines zones ne sont pas systématiquement acquises. En 3.14, on observe l'apparition de nuages qui se traduisent par l'apparition sur une seule image de couples nuage-ombre qui ont des formes non géométriques. Notons dans les séquences a), b) et d) des figures 3.3 et 3.4 l'existence de zones non géométriques blanches que l'on aurait pu prendre pour des nuages si ce n'était l'absence d'ombre si l'on ne disposait que d'une image. La séquence dont on dispose nous permet d'observer la répétitivité de ces régions, et donc d'écarter l'hypothèse du nuage. D'après [Belward and Valenzuela, 1991], des variations de tons constantes dans le temps sont vraisemblablement liées à la nature du sol alors que les variations tonales liées non constantes dans le temps sont vraisemblablement liées à l'humidité du sol. Ces zones sont effectivement constantes dans le temps sur une petite période ce qui nous permet de les distinguer des nuages, cependant, elles ne sont pas tout le temps présentes. De plus, on les voit même lorsqu'il y a des cultures. Ceci nous amène à penser qu'il s'agit probablement de zones plus sèches.

La STIS présente donc une certaine stabilité géométrique. Hormis quelques changements brutaux tels que la division d'une parcelle de terre nue liée à la croissance de différentes cultures ou à un labour ou une récolte en plusieurs temps, les objets qui constituent la scène gardent une même forme. La radiométrie, en revanche peut varier beaucoup.

Pour des STIS composées d'images mieux résolues comme celles qu'acquerront les satellites Pleïades, d'autres phénomènes interviendront. La stabilité géométrique et la variabilité radiométrique restent néanmoins des propriétés qui caractérisent ces STIS en raison de la nature de la scène observée et du sous-échantillonnage temporel et constituent donc les deux hypothèses de base de ce travail. Ceci sera discuté plus en profondeur au ch.4.

Nous avons choisi d'exploiter cette stabilité géométrique en extrayant ces objets dont la forme est généralement conservée au cours du temps de façon à pouvoir ensuite analyser l'évolution de chaque objet.

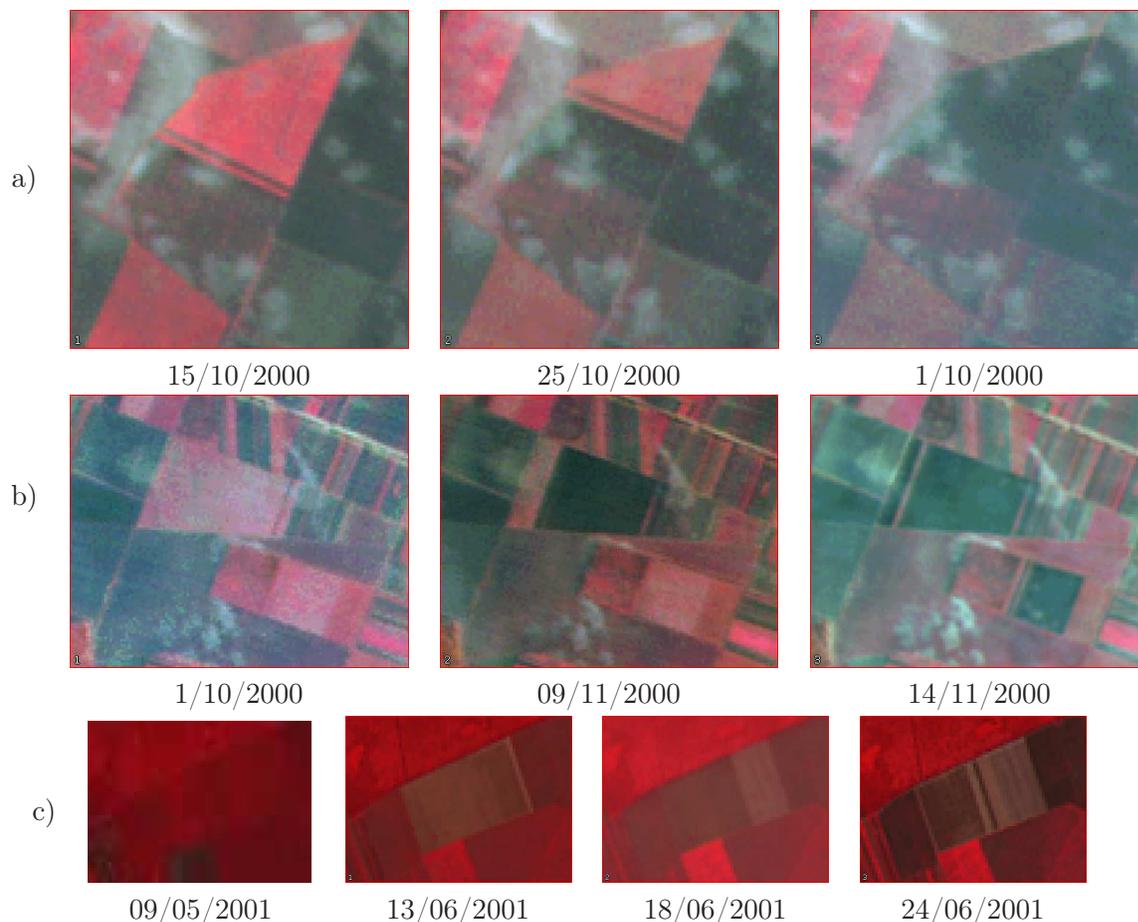


FIGURE 3.3 – Récoltes. Dans la séquence d’images a), on remarque par ailleurs que les zones ont été moissonnées en deux étapes : la première fois, entre le 15 octobre et le 25 octobre, et la deuxième fois, entre le 25 octobre, et le 31 octobre. On sait par ailleurs que dans cette zone géographique, parmi les cultures les plus pratiquées, le maïs se récolte entre le 10 octobre et le 10 novembre, et, les betteraves entre le 25 septembre et le 20 novembre. Il s’agit donc probablement d’un champ de betteraves ou de maïs. L’intensité du rouge de la zone avant récolte indique qu’il s’agit probablement de betteraves car le maïs est généralement plus sec au moment de la récolte. On pourrait définitivement lever l’ambiguïté en utilisant des images de mars-avril 2000 étant donné que les betteraves lèvent fin mars alors que le maïs lève courant avril, mais pour cet exemple, il s’agit des trois premières images de la séquence. Dans la séquence d’images c) (qui se poursuit sur j)), on observe un blanchissement par parties d’une zone. D’après [Belward and Valenzuela, 1991], des zones claires peuvent correspondre à de la terre sèche ou calcaire. Après récolte, l’activité végétale réduit rapidement et la terre sèche. On peut donc supposer que ce blanchissement par parties d’une parcelle correspond en fait à l’assèchement de la terre suite à la récolte en plusieurs fois. Seuls le colza et l’orge se récoltent à cette période. Etant donné que cette zone n’a pas subi de blanchissement traduisant une floraison (voir figure 3.8), on peut écarter l’hypothèse du colza et supposer qu’il s’agit d’une culture d’orge.

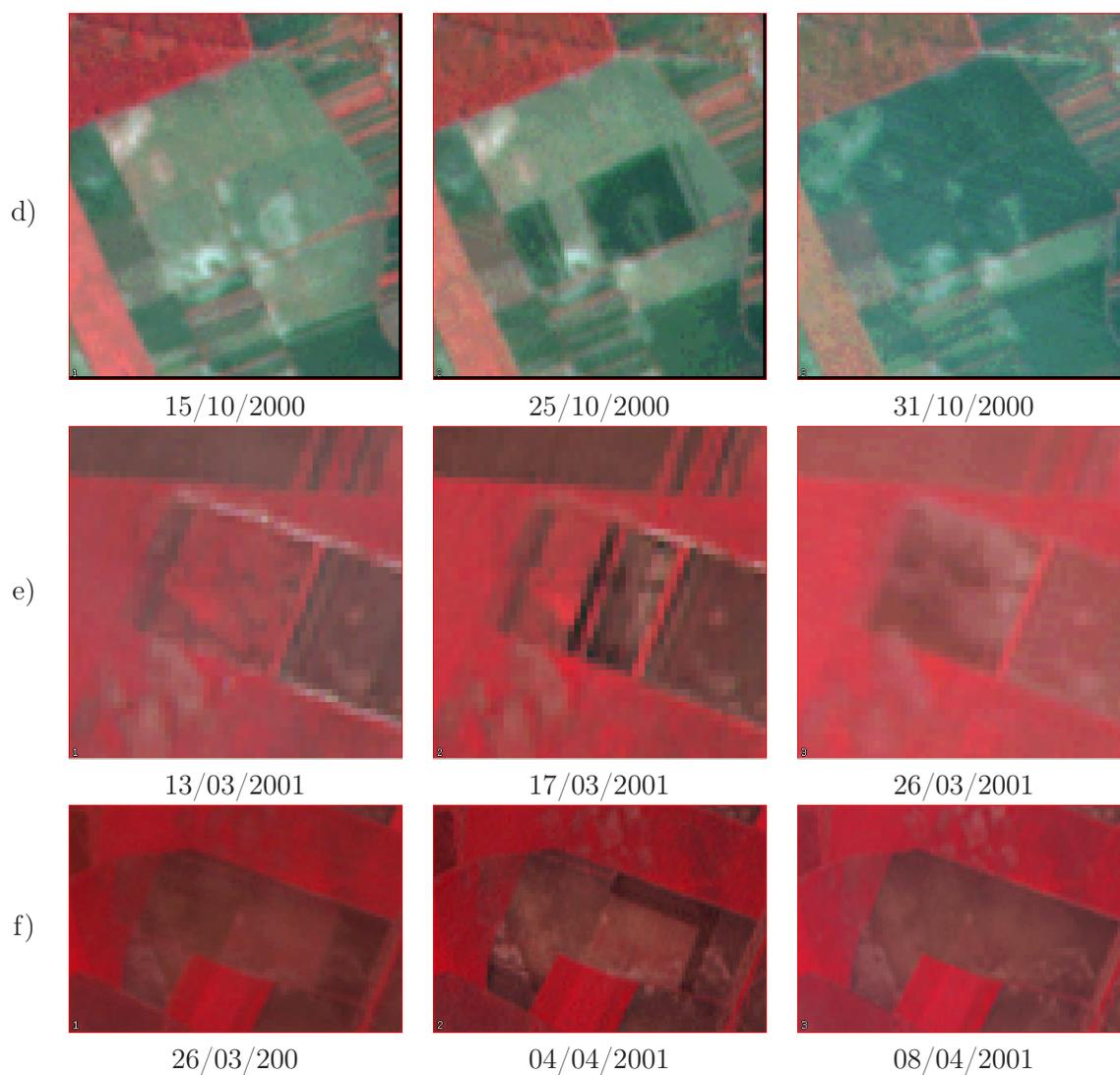


FIGURE 3.4 – Labour. Dans l'exemple d) de la figure 3.4, la brutalité du changement indique ici encore qu'il s'agit d'un phénomène lié à une intervention humaine. La transition radiométrique allant du clair vers le sombre, on peut supposer que la zone sombre correspond au sol nu d'où l'interprétation par un labour. Le cas e) ressemble aux cas de récolte a), b), et c), si ce n'est qu'on ne récolte pas en mars. Il s'agit donc ici encore d'un labour : la zone rouge précédant l'apparition de la zone sombre correspondant à des herbes folles. Il en est de même pour l'exemple f).

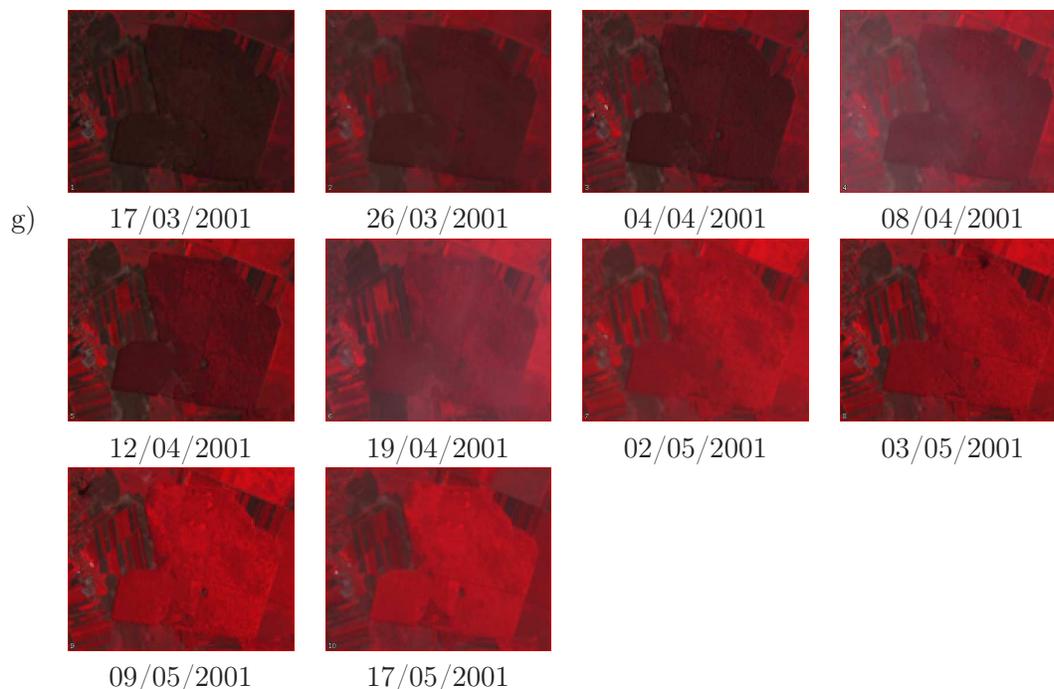


FIGURE 3.5 – Forêt. Cette séquence d’imagettes illustre l’évolution d’une forêt de feuillus. En effet, la taille, et la forme “complexe” –Les parcelles sont généralement de forme rectangulaire, éventuellement trapézoïdales, mais présentent rarement de concavité. Les forêts, en revanche peuvent avoir des formes géométriques ou pas et peuvent présenter des concavités.– indique qu’il s’agit d’une forêt, et la couleur sombre de la zone en hiver indique qu’il s’agit d’arbres à feuilles caduques. On peut observer le rouge s’intensifier progressivement de mars jusqu’en mai où la réflectance maximale en infra-rouge est atteinte.

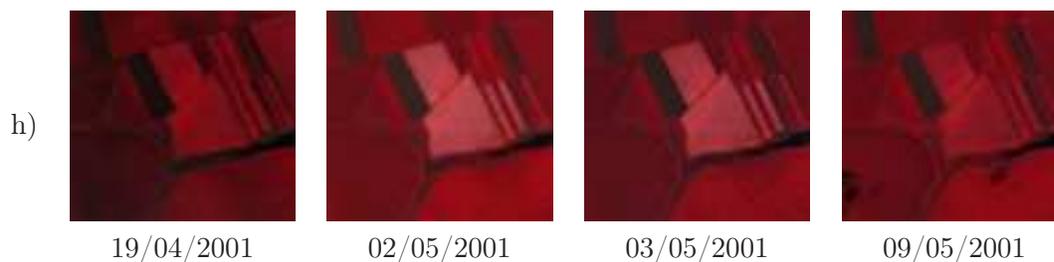


FIGURE 3.6 – Floraison. On observe un blanchissement de zones végétalement actives. La zone blanchie continue de réfléchir dans l’infra rouge étant donnée la teinte rosée de la zone ce qui indique que la végétation est toujours présente. Un sol sec serait moins rouge. D’après [Belward and Valenzuela, 1991], la floraison se traduit par un blanchissement de la zone car les fleurs réfléchissent dans le visible, et cachent les parties vertes qui sont celles qui réfléchissent dans l’infra rouge. Il s’agit donc vraisemblablement d’un champ de colza.



FIGURE 3.7 – Neige. On observe l'apparition brutale d'un contour blanc autour des objets, et ce contour blanc disparaît partiellement sur l'image suivante. Il s'agit probablement de neige. Il aurait neigé entre le 22 février et le 8 mars, et la neige aurait fondu sur les champs, et serait restée particulièrement sur les bords de route et sur les chemins.

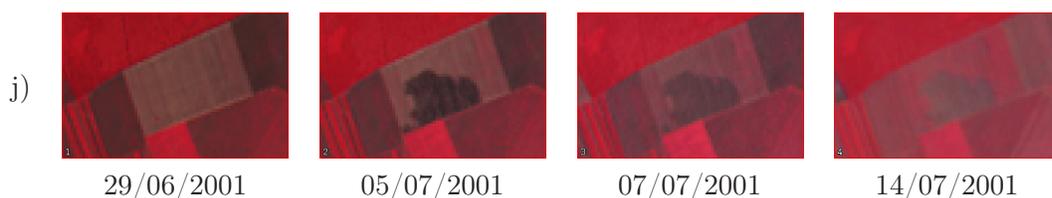


FIGURE 3.8 – Brûlis. On voit apparaître une zone naturelle sombre. L'assombrissement est souvent lié à de l'humidité dans la terre. Cependant étant donnée la date, il paraît peu probable que la terre s'humidifie de façon naturelle sur cette zone. On sait par ailleurs que la culture sur brûlis se pratique dans cette zone; il s'agit donc plus probablement d'un brûlis.

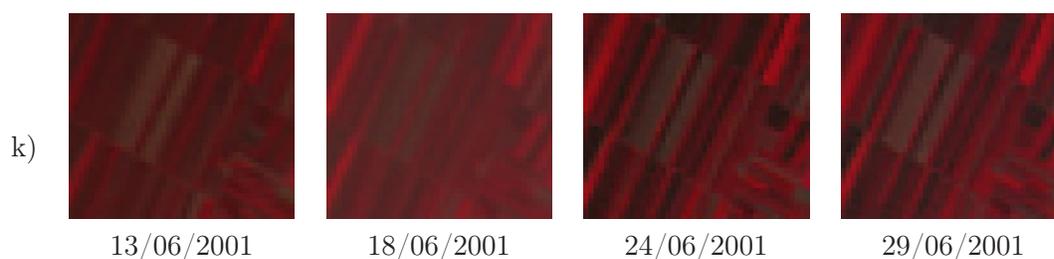


FIGURE 3.9 – Différentiation des cultures. On observe la différenciation d'une zone labourée en différentes parcelles couvertes par des cultures de nature différente et dont la radiométrie évolue indépendamment. On voit ici comment des changements radiométriques locaux provoquent des changements topologiques.

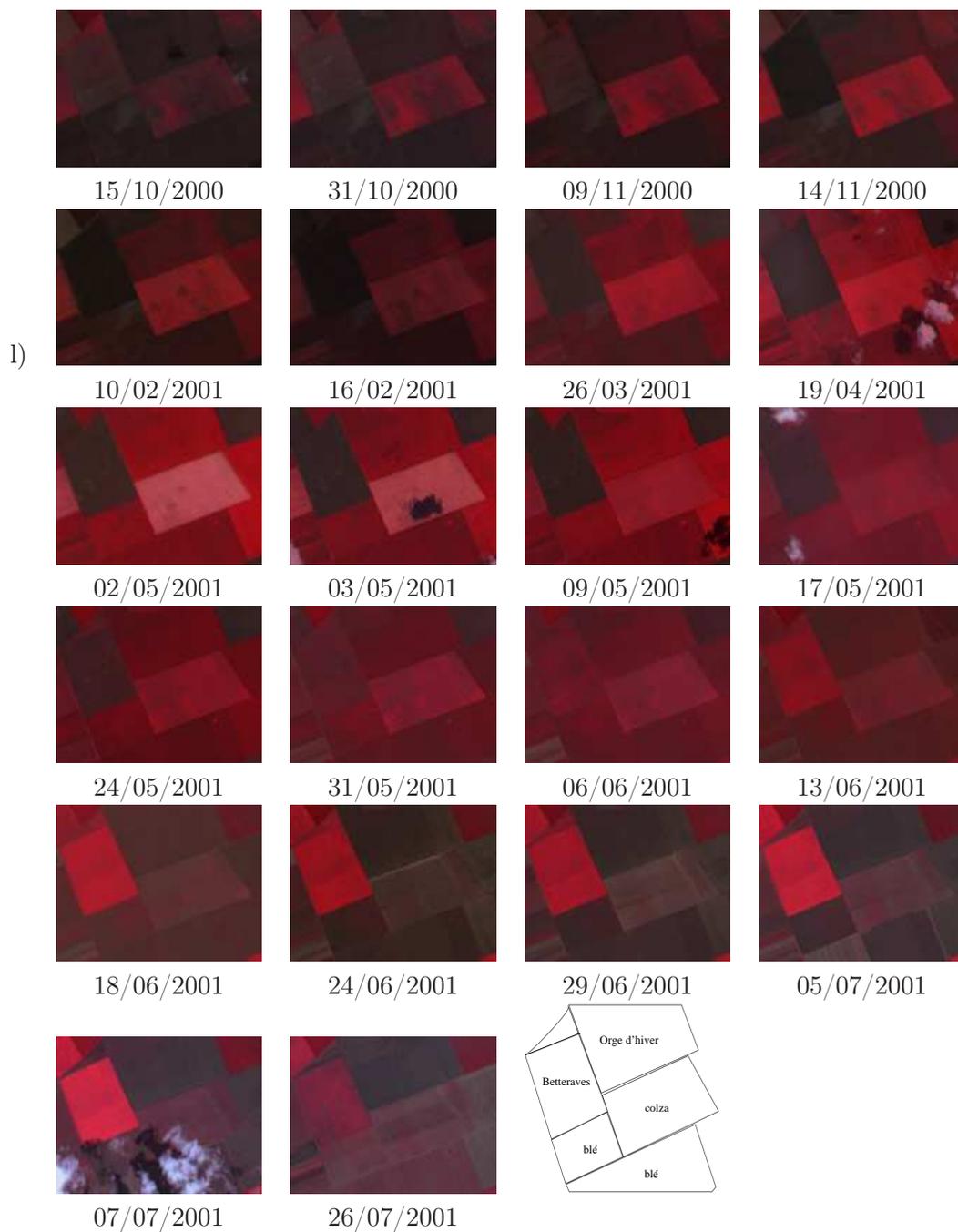


FIGURE 3.10 – On identifie facilement le colza pour la floraison en mai, le blé, qui lève en mars et est récolté fin juin, la betterave qui lève en mai, et est récoltée après la fin de la STIS (17 juillet), et l'orge d'hiver qui lève fin octobre début novembre, est récolté fin juin (Normalement la récolte de l'orge d'hiver a plutôt lieu début juillet. Peut-être s'agit-il d'une autre culture.).

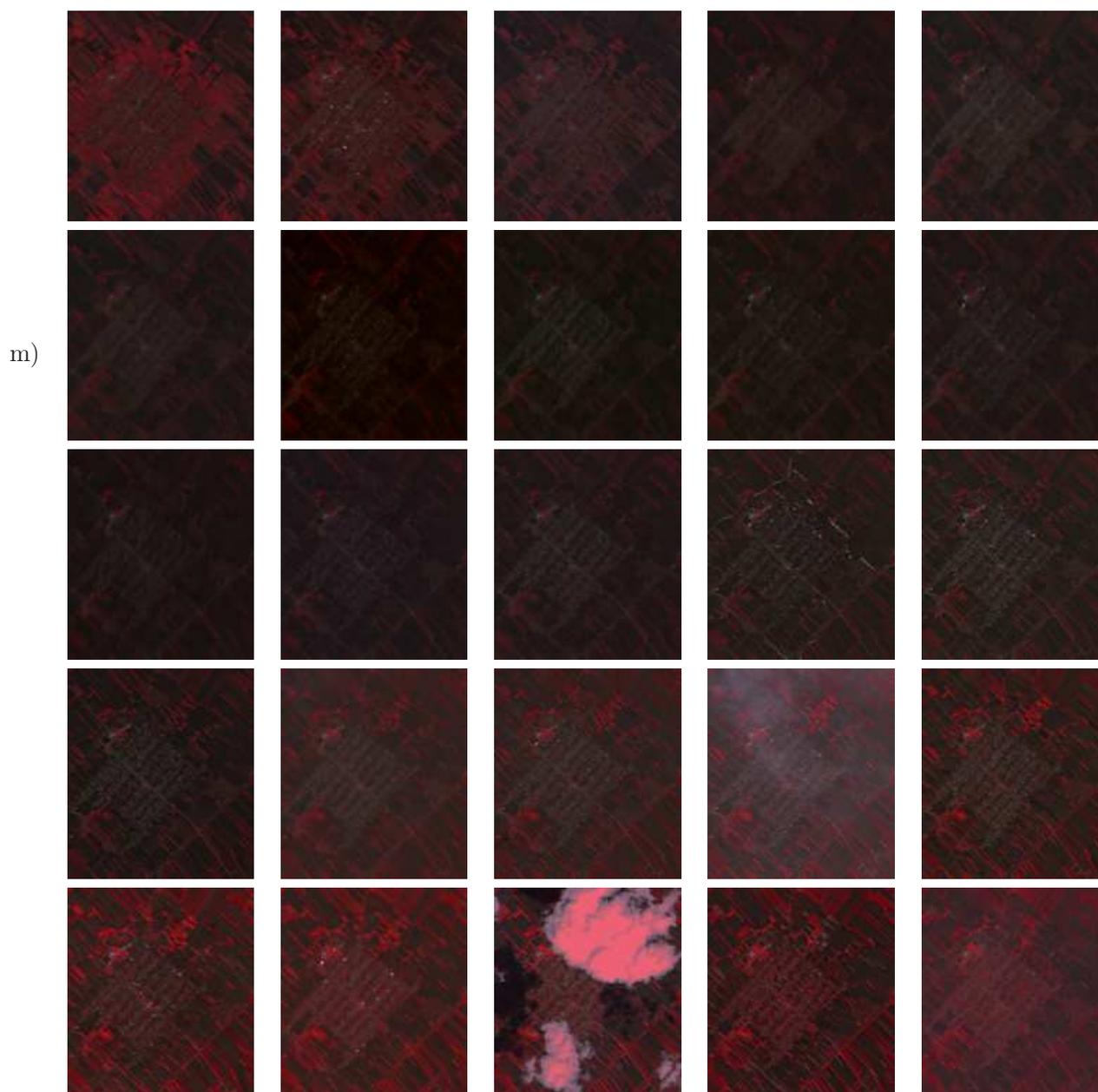


FIGURE 3.11 – Ville. Certains points brillants apparaissent et disparaissent dans le temps. Un zoom sur l'un de ces points est présenté figure 3.12. Ces points correspondent à des réflexions spéculaires de surfaces lisses et dépendent donc de l'orientation de la prise de vue et de la position du soleil ce qui explique les scintillements.

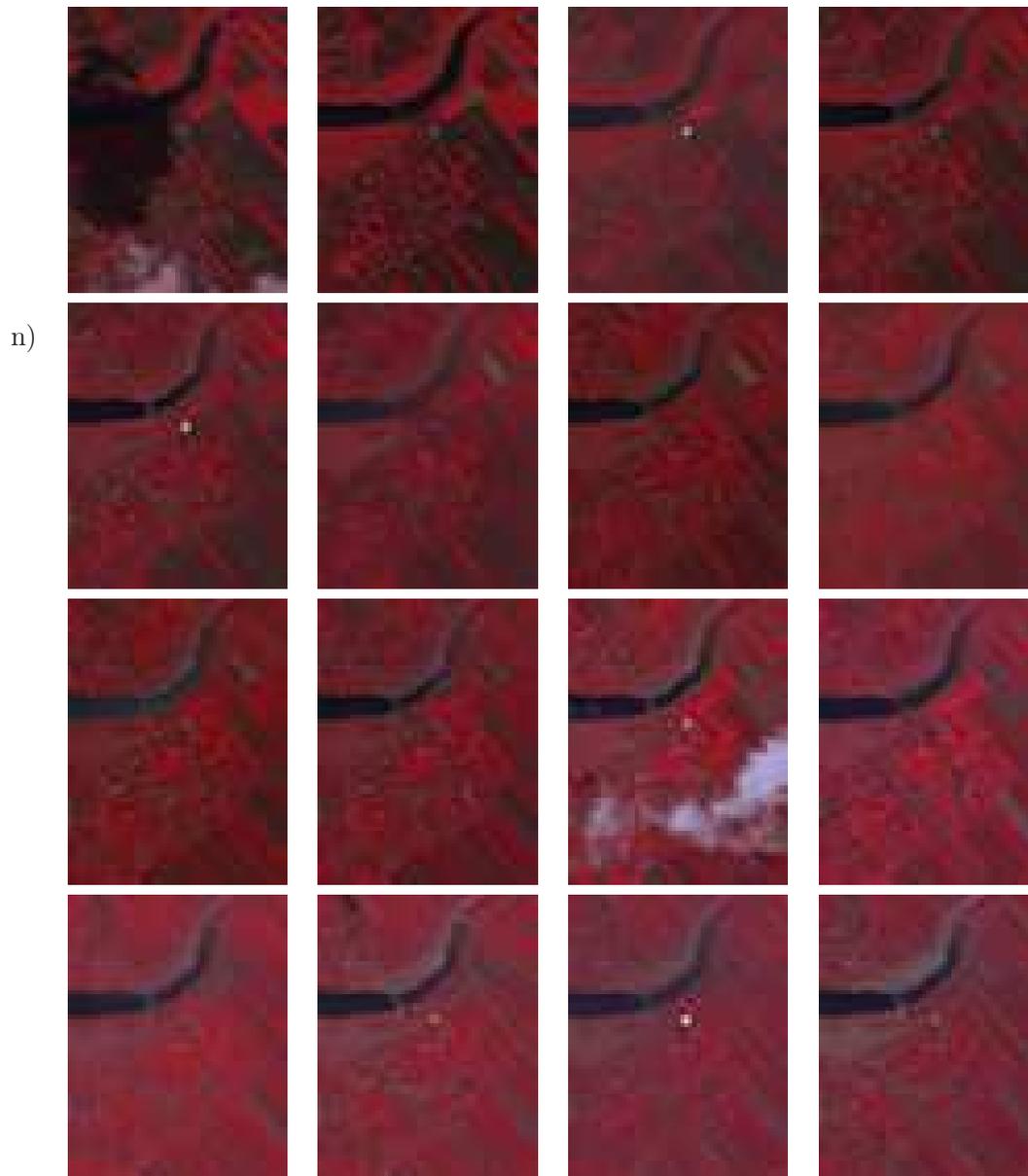


FIGURE 3.12 – Zoom sur la ville de la figure 3.11.

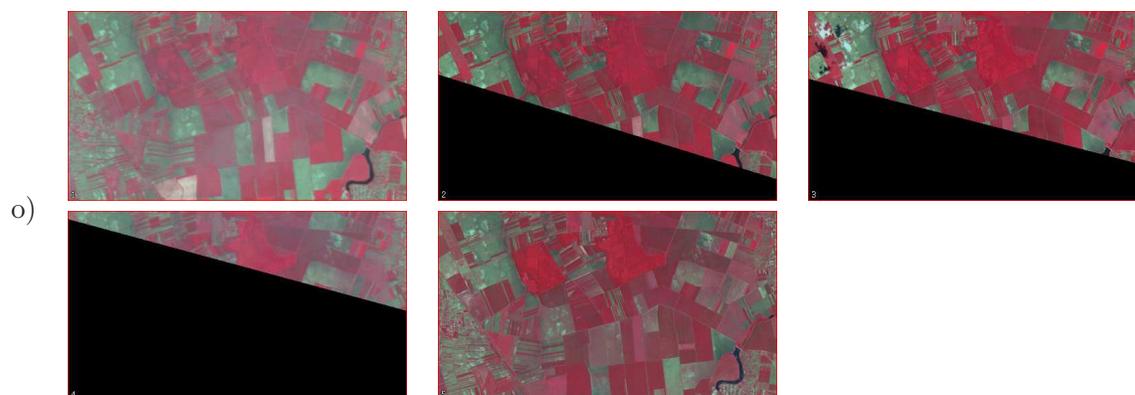


FIGURE 3.13 – Support spatial variable. En raison des variations d’angle lors de l’acquisition, en certains instants, certaines zones ne sont pas imagées.

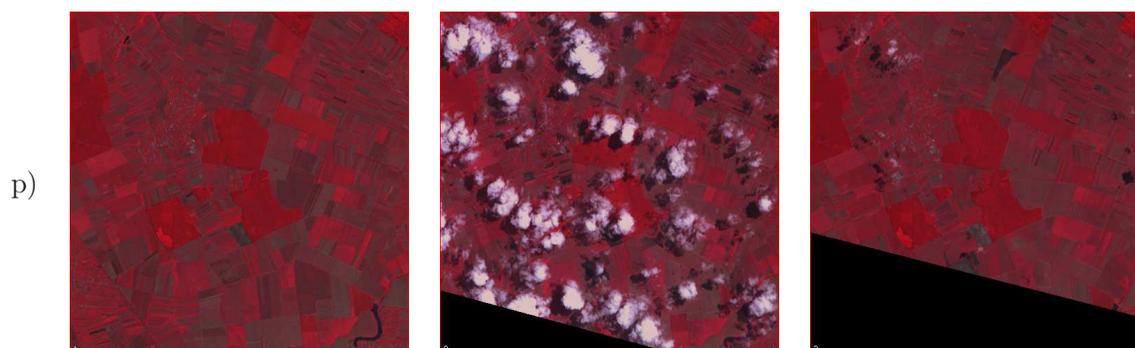


FIGURE 3.14 – Nuages.

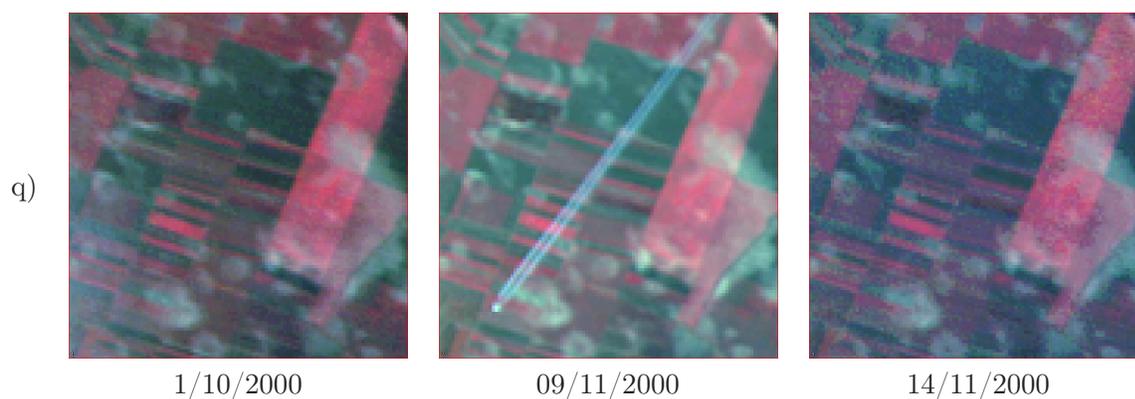


FIGURE 3.15 – Traînée d’avion. L’apparition d’un objet sur une seule image sans conséquence sur les images suivantes indique qu’il s’agit probablement d’un objet entre le sol et le capteur. Dans ce cas, la forme rectiligne indique qu’il s’agit de la traînée d’un avion.

3.2 Approche adoptée

Ce nouveau type de données nous permet d'identifier des phénomènes spatio-temporels. Par exemple identifier une récolte, un labour, l'apparition de nuages, éventuellement inférer ce qui se trouve dessous, identifier l'évolution radiométrique d'un certain type de culture, trouver des fréquences temporelles d'évolution, etc. Afin de pouvoir extraire ce type d'information de façon automatique, il faut modéliser la séquence. Le modèle ne doit pas atteindre le niveau de la sémantique, mais doit structurer l'information de façon qu'on puisse effectuer des requêtes à l'aide d'exemples, ou de modélisations d'un phénomène particulier, et retrouver des réalisations de phénomènes spatio-temporels similaires.

Nous avons vu dans la partie I différentes modélisations d'images, et de séquences d'images. Cependant, aucune n'est adaptée à notre problématique. C'est pourquoi, nous proposons dans le ch.4 un modèle original adapté aux STIS.

Troisième partie
Cadre théorique

Chapitre 4

Approche de l'estimation du modèle de STIS

Nous avons vu au §A.2.4.1 que les modèles existants de scène dynamique par satellite supposent une géométrie invariante dans le temps, ainsi qu'une radiométrie variable. Les résolutions spatiales et temporelles des systèmes d'acquisition vont en s'améliorant, et on est maintenant capable d'observer des changements dans la géométrie de la scène. Un modèle plus complet de la description de la scène serait donc bienvenu. Nous avons vu au §A.2.3.3 des modèles permettant de décrire des mouvements et des déformations. Afin de prendre en compte ces changements caractéristiques d'une haute résolution spatiale et d'une fréquence d'acquisition élevée, nous proposons un modèle mixte.

4.1 Scène statique

Une scène statique est composée d'objets spatiaux (SO) définis par une région spatiale et un vecteur de caractéristiques. Ces deux attributs définissent un objet de façon unique et on considère donc un identifiant pour chacun d'eux. Les caractéristiques d'un SO peuvent être radiométriques ou texturales, et ces caractéristiques sont homogènes à l'intérieur de l'SO, et changent lorsqu'on passe la frontière.

4.2 Vers un modèle de scène dynamique de télédétection à haute résolution spatiale et temporelle

Dans une scène dynamique, les objets changent : les véhicules se déplacent, apparaissent et disparaissent de la scène, un feu de forêt s'étend, un nuage survole une zone, la radiométrie d'une culture varie, une zone agricole se divise en plusieurs parcelles au printemps, et ces parcelles peuvent fusionner lors de leur croissance, lors d'une récolte ou lors d'un labour. Les changements correspondant à ces phénomènes sont respectivement des déplacements rigides, des apparitions et disparitions, des déformations, des superpositions, des évolutions de caractéristiques (radiométriques), des divisions, des fusions. Plusieurs remarques sont à faire par rapport à ces changements : d'une part, certains d'entre eux impliquent de considérer de nouveaux attributs pour un objet spatio-temporel (STO). Ainsi, l'apparition et la disparition d'objets supposent que les STO soient aussi caractérisés par un intervalle temporel, et la superposition d'objets nécessite de considérer différentes profondeurs pour différents objets. D'autre part, ces changements agissent sur tous les attributs d'un objet.

Or pour parler d'STO, il faut une homogénéité spatiale et temporelle d'un des attributs. Nous allons donc considérer différents types d'objets en fonction des changements qu'ils peuvent subir.

4.2.1 Profondeur des objets

Nous avons vu au §A.2.3.3 que les déplacements et les déformations rendent la prise en compte de la troisième dimension spatiale nécessaire pour une description cohérente. En effet, les objets se déplaçant ou se déformant couvrent et découvrent des zones se situant en retrait. Nous allons donc considérer plusieurs plans dans la scène dynamique. Le sol ou sa couverture correspond à l'arrière plan, et les différents objets en mouvement spatial, à différentes profondeurs. La profondeur de l'arrière-plan est arbitrairement fixée à zéro. Cette dernière peut néanmoins changer. En effet, l'arrière-plan se distinguera des autres plans par les changements que peuvent subir les objets qui les composent et, pour une modélisation complète, il nous faut alors permettre aux objets de l'arrière-plan de passer au premier plan pour subir les changements qui y sont permis.

4.2.2 Définition d'un objet spatio-temporel (OST)

La scène dynamique que nous considérons est composée d'objets spatio-temporels (STO) ayant une étendue spatiale et temporelle et de sorte que la projection d'un STO sur un instant soit un SO temporellement localisé. Un STO a donc quatre attributs et un identifiant, i : la région spatiale changeante, R_i^c , l'étendue temporelle, $[[t_i^1, t_i^2]]$, le vecteur de caractéristiques changeant, Ψ_i^c , et la profondeur changeante, d_i .

$$O_i = (R_i^c, \Psi_i^c, d_i, [[t_i^1, t_i^2]]) \quad (4.1)$$

Nous avons vu au §2.1 que l'on peut distinguer deux types de changements : les événements, et les processus. Les processus sont caractérisés par une durée temporelle alors que les événements sont des changements d'état. Les événements s'appliquent aux caractéristiques, aux régions, à la profondeur, et à l'étendue temporelle. On parle respectivement d'événement de caractéristiques, \mathcal{E}^Ψ , d'événement spatial, \mathcal{E}^R , d'événement de profondeur, \mathcal{E}^d , et d'événement d'existence, $\mathcal{E}^{\Delta t}$. Les processus peuvent s'appliquer aux régions et aux caractéristiques. On les appelle respectivement processus spatial, \mathbb{P}^R , et processus de caractéristiques, \mathbb{P}^Ψ . En développant les différents changements des attributs de l'STO, l'eq.4.1 devient :

$$O_i = \left((R_i^0, \{\mathbb{P}_i^R\}, \{\mathcal{E}_i^R\}), (\Psi_i^0, \{\mathbb{P}_i^\Psi\}, \{\mathcal{E}_i^\Psi\}), (d_i^0, \{\mathcal{E}_i^d\}), ([[t_i^1, t_i^2]], \mathcal{E}^{\Delta t}) \right) \quad (4.2)$$

4.2.3 Processus des OST

On distingue deux types de processus spatiaux en raison de leur différence de modélisation : la déformation et le mouvement rigide. Les processus spatiaux n'agissent que sur les objets des premiers plans. Ainsi, un véhicule se déplaçant aura une profondeur non nulle.

Les caractéristiques des objets peuvent aussi évoluer. Les processus de caractéristiques correspondent à des modèles typiques d'évolution. Par exemple, l'évolution selon la courbe de Fischer est un processus typique des cultures pour la caractéristique indice de végétation par différence normalisée (Normalized Difference Vegetation Index, NDVI). Ces processus de caractéristiques induisent un autre changement de nature structurelle. En effet, considérons deux SO spatialement voisins SO_i^t et SO_j^t obtenus par projection au temps t des

STO O_i et O_j . A cet instant, les deux STO ont des caractéristiques distinctes. Cependant, au cours de leur évolution, il peut arriver un instant t' en lequel les instances des STO ont des caractéristiques similaires. Dans ce cas, afin d'assurer que la projection de l'STO en t' soit un SO, il faut ne considérer plus qu'un des deux objets. Réciproquement, un comportement particulier des caractéristiques d'une partie d'un STO aboutira en une division. Nous voyons ici que la modélisation d'une scène en STO homogènes spatialement selon des caractéristiques, et dont ces caractéristiques évoluent dans le temps implique des changements topologiques qui sont des événements spatiaux. Détaillons maintenant les différents événements.

4.2.4 Événements des OST

Les événements peuvent s'appliquer aux régions, aux caractéristiques, à la profondeur, et à l'étendue temporelle. On parle alors respectivement d'*événement spatial*, d'*événement de caractéristiques*, d'*événement de profondeur*, et d'événement d'existence.

Il existe deux événements d'existence : la *création* et la *cessation*. Pour l'arrière-plan, ces événements sont exclusivement liés à l'étendue temporelle bornée de la scène et n'existent donc qu'au début et à la fin de la séquence. Pour les objets des premiers plans, en revanche, des créations et cessations peuvent avoir lieu entre les instants de début et de fin de la scène. Par exemple, un bateau sur une rivière apparaîtra d'un bord spatial de la scène là où commence la rivière à un certain instant, et disparaîtra à l'autre bord. La description dans notre modèle de ce phénomène est donc la création à un instant $t \in [0, T - 1]$ d'un STO de profondeur 1, de caractéristique statique et dont la région est en mouvement rigide pendant une durée d , et la cessation de l'objet en $t + d$.

Les événements spatiaux sont la *fusion* et la *division*. Ces événements n'agissent que sur les objets de l'arrière-plan. En effet, les événements spatiaux sont provoqués par les évolutions des caractéristiques, restreindre les événements spatiaux aux objets de l'arrière-plan revient à restreindre les évolutions de caractéristiques aux objets de l'arrière plan. La description d'un glissement de terrain est une création d'un STO de profondeur non nulle en déformation.

Nous avons vu que les processus des caractéristiques des objets étaient liés à une évolution typique d'une couverture de sol par exemple une culture de blé, de colza, de maïs, ou encore une forêt de mélèzes, d'épicéas, de pins, de peupliers, etc. La couverture du sol d'une même zone spatiale peut donc changer dans le temps : typiquement d'une année sur l'autre selon les rotations de cultures. Il s'agit alors d'événements de caractéristiques. Notons que cet événement n'est pas nécessairement accompagné d'un événement spatial. En effet, on peut considérer qu'une culture dure depuis son semis jusqu'à sa récolte. La récolte peut être liée à un événement spatial, le semis, en revanche ne l'est pas.

Les STO sont aussi caractérisés par leur profondeur. Or cette dernière peut changer en fonction du comportement de l'objet. Ainsi, une rivière en crue change de profondeur au moment où elle dépasse son lit. En effet, la région associée est alors en déformation, et seuls les objets des premiers plans peuvent subir ce type de transformation. On considère donc aussi que la profondeur des STO peuvent être modifiés par des événements.

4.2.5 OST et sémantique

Tout au long de l'exposé, nous donnons des exemples concrets de phénomènes à rattacher à ces STO. Ces exemples correspondent à une interprétation sémantique des STO qui n'est pas dans le modèle. Le modèle que nous proposons est constitué d'STO dont

les attributs subissent des événements. Un événement définit un changement d'état ce qui peut se traduire sémantiquement par la cessation d'(un) objet(s) et la création d'(un) autre(s). Ainsi, on peut décrire le phénomène suivant : *La zone agricole définie par la région spatiale R_i s'est divisée en les parcelles définies par les régions spatiales R_j , R_k , et R_l .* Cette description suggère que la zone agricole cesse d'exister au moment où les parcelles apparaissent. En conférant une dimension sémantique aux STO, on serait alors tenté de supposer des événements d'identité pour les objets de l'arrière-plan aussi, et des STO avec une étendue temporelle entre deux événements spatiaux. On peut cependant aussi décrire le phénomène : *La culture de maïs située en R_i a été semée le 10 avril, a levé le 12, est arrivée à maturation le 15 août, et a été récoltée le 17.* Dans ce cas, l'objet considéré traverse les événements spatiaux et les événements d'identité coïncident avec les événements des caractéristiques. Les deux descriptions correspondent à deux points de vue différents et à des sémantiques différentes. Nous allons d'ailleurs nous intéresser à ces deux approches. Cependant, afin que notre modèle reste valide pour ces deux analyses, nous repoussons les décisions d'identité à une étape ultérieure. Ce point de vue est aussi celui adopté dans [Raza et al., 2000] via le concept de propriété-clé. Ainsi, dans le premier cas, l'intérêt porte sur la spatialité des objets. La propriété-clé est alors la région spatiale des STO et elle permet de définir des STO sémantiques dans un contexte spatial. Dans le second cas, la propriété clé est la caractéristique, et elle permet de mettre en évidence des STO sémantiques dans un contexte de caractéristiques.

De même, pour les événements de profondeur, plusieurs descriptions sémantiques sont envisageables : on peut considérer la rivière en crue comme une entité particulière différente de la rivière. En effet, cet objet a des caractéristiques différentes, sa région évolue via un processus, et sa profondeur a changé. Mais on peut aussi considérer qu'il s'agit d'une variation du même objet. Cette décision sera prise lors de l'interprétation.

4.2.6 Conclusion

Nous avons proposé un modèle de scène de télédétection à haute résolution spatiale et temporelle permettant de prendre en compte une grande variété de changements. Ce modèle est un modèle mixte entre les modèles de scène de vidéo multimédia et de scènes de télédétection à résolution grossière. Il considère plusieurs plans de façon à gérer les occlusions et désocclusions de l'arrière-plan liées aux déplacements et déformation comme c'est classiquement fait en vidéo. Les objets des différents plans sont caractérisés par leur profondeur changeante, par leur région spatiale changeante, par leurs caractéristiques changeantes, et par leur étendue temporelle changeante. Les changements de chacun de ces attributs peuvent être de deux formes : événement ou processus. Les processus spatiaux et événements d'existence sont restreints aux STO des premiers plans, et les événements spatiaux et processus de caractéristiques sont restreints aux objets de l'arrière plan. Comme en vidéo, les objets des premiers plans (en se déplaçant ou se déformant) ont des caractéristiques statiques, et peuvent apparaître et disparaître. Les caractéristiques des objets de l'arrière-plan en revanche peuvent évoluer sur une durée suivant un processus ou de façon brutale via un événement. Les objets d'arrière-plan peuvent passer au premier plan via un événement de profondeur, et réciproquement, des objets des premiers plans dont le processus de région cesse, passent à l'arrière-plan.

4.3 Hypothèses simplificatrices pour l'estimation d'un modèle de STIS

Nous avons décrit un modèle de scène. Voyons maintenant comment estimer ce modèle dans une STIS.

4.3.1 Complexité de l'estimation

Décomposer une STIS selon ce modèle est une tâche complexe. En effet, comme nous l'avons vu au §A.2.3.3, la détermination des STO nécessite non seulement une homogénéité spatiale, mais aussi une homogénéité temporelle. Dans notre modèle, l'homogénéité temporelle n'est pas la même pour les différents objets. Ainsi, les objets de l'arrière-plan présentent une homogénéité temporelle de forme, ¹, et les objets des premiers plans présentent une homogénéité temporelle radiométrique, et de forme dans le cas des déplacements rigides. Ces homogénéités sont donc de natures différentes pour les différents types d'objets, et ces différents types d'objets sont justement définis par ces différences. Cette double dépendance rend leur estimation difficile.

4.3.2 Sous-échantillonnage temporel

Pour la séquence ADAM, la période d'échantillonnage varie entre un jour et un mois. Pour RapidEye, la période d'échantillonnage est de un jour, et pour Pléiades, on attend une période du même ordre. Les fréquences d'acquisitions des satellites à haute résolution spatiale et temporelle actuels ne permettent donc pas d'estimer des déplacements et déformations trop rapides. En effet, d'après le théorème de Shannon, une fréquence d'échantillonnage journalière permet de caractériser des phénomènes dont la période minimale est supérieure à deux jours. Or la plupart des phénomènes d'intérêt aux résolutions actuelles, par exemple le déplacement d'un véhicule, ont des périodes maximales bien inférieures à deux jours. Ces variations géométriques d'objets plus petits étaient moyennées dans des STIS-CR dont la résolution est de l'ordre du km et ce point n'était donc pas gênant. En améliorant la résolution spatiale, les objets visibles sont plus petits, et leur géométrie commence à varier significativement.

Les processus spatiaux observables pour une fréquence d'acquisition journalière sont par exemple les variations de hauteur d'un lac ou d'une rivière asséchée, ² (déformation), des variations de l'humidité du sol (déformation), le déplacement de l'ombre des arbres en raison de la hauteur du soleil.

4.3.3 Hypothèse de scène arrière plan

Afin d'extraire les objets, nous allons donc supposer que la scène ne contient que des objets de l'arrière-plan, et permettre de petits déplacements et déformations que nous regrouperons dans une incertitude géométrique appelée *bruit géométrique*. La profondeur des objets, pourra ensuite être corrigée en analysant les STO obtenus, et leurs changements.

1. cette homogénéité concerne des groupes d'objets en raison des événements spatiaux. Ce point sera détaillé au ch.9.

2. les variations de largeur d'une rivière dans son lit restent relativement lentes, en revanche, si elle sort de son lit, la déformation est trop rapide pour être décrite de façon continue avec une acquisition journalière



FIGURE 4.1 – Exemple de phénomène dont la description sous forme de processus spatial n'est pas possible en raison de l'échantillonnage de la STIS, et dont la description par événements spatiaux nécessite l'introduction d'un événement hybride : la combinaison de fusions et de divisions.

La scène peut toutefois présenter occasionnellement des déformations ou des déplacements importants. Ainsi, la figure 4.1 présente l'exemple d'un labour qui s'est effectué en plusieurs fois. Le phénomène correspondant est l'apparition d'une région en expansion, cependant, la fréquence d'acquisition ne nous permet pas de le décrire ainsi. L'amplitude de cette déformation étant trop importante, on ne peut pas supposer qu'il s'agit d'un bruit géométrique. Afin d'avoir une description géométrique complète de la scène, nous proposons donc d'introduire un autre changement compensant les erreurs liées à l'hypothèse d'une scène arrière-plan. Ce changement sera donc aussi décrit par un événement spatial : la combinaison de fusions et de divisions ou simplement *combinaison*.

Nous obtenons ainsi une description géométrique complète de la scène. En effet, on peut toujours passer d'une partition à une autre en effectuant des fusions et des divisions dans chacune des partitions. Ces trois types d'événements topologiques (la fusion, la division, et la combinaison) permettent donc une description géométrique de toute STIS. Ainsi, lors d'une déformation d'un état initial à un état final, si l'on dispose d'acquisitions intermédiaires, la déformation pourra être décrite par une succession de combinaisons.

Un compromis devra donc être trouvé selon l'amplitude du processus spatial entre une description par combinaison ou par un bruit géométrique.

4.3.4 Conclusion

L'estimation des STO, de leur changements, et de leur profondeur est trop difficile sans *a priori*, et étant donné que la fréquence d'acquisition des satellites actuels ne permet pas de caractériser des déplacements ou déformations trop importants, nous proposons pour la détermination des SO, instances des STO, de considérer que la scène ne contient que des objets de l'arrière-plan. Par ailleurs, afin de prendre en compte de faibles déformations ou déplacements, nous considérons un bruit géométrique.

4.4 Approche de l'analyse de l'arrière plan

Le modèle que nous avons proposé est constitué d'STO dont la région spatiale n'est modifiée qu'à travers des événements spatiaux et dont les caractéristiques changent via des processus et des événements. Etant donné que les objets de l'arrière-plan ont tous une profondeur nulle, et que leur durée de vie est la durée de la STIS, nous omettrons par la suite ces attributs. Un STO de l'arrière-plan est donc caractérisé par :

- Une région spatiale, et les événements associés : $(R_i, \{\mathcal{E}_i^R\})$

- Un vecteur de caractéristiques, et les processus et événements associés : $(\Psi_i, \{\mathbb{P}_i^\Psi\}, \{\mathcal{E}_i^\Psi\})$

Afin de déterminer les objets, nous avons supposé que la scène ne contient que des objets de l'arrière-plan. Les déformations et déplacements sont néanmoins pris en compte, par un bruit géométrique s'ils sont de faible amplitude, et par une combinaison, sinon. Un compromis devra donc être trouvé pour choisir la meilleure description.

Le bruit géométrique considéré permettra par ailleurs de prendre en compte les éventuelles erreurs d'intercalibration liées à d'éventuels changements de conditions atmosphériques, d'éclairage, d'angle de visée entre les acquisitions, et de capteur. Or, nous avons vu qu'une des propriétés des STO est que les caractéristiques de leurs différentes instances doivent être homogènes spatialement. Afin de satisfaire cette propriété, nous considérons les différentes instances, que nous chaînons en utilisant des critères de mise en correspondance de forme. Afin de prendre en compte aussi les changements topologiques des régions des STO, nous adoptons une structure de graphe : le graphe d'adjacence temporelle des instances d'STO.

4.4.1 Description de la scène par un graphe

On considère une description complète de la scène. Autrement dit, tout point de l'espace spatio-temporel discrétisé appartient à une instance d'STO, c'est-à-dire un SO. Ces SO sont alors reliés par des relations d'adjacence temporelle.

Les STIS ne présentent pas les STO de façon explicite, et notre objectif est de les extraire. Nous considérons donc une structure intermédiaire dont les STO se déduisent. Il s'agit du graphe d'adjacence temporelle des objets spatiaux, SOTAG (Spatial Object Adjacency Graph).

Soient $(\mathcal{P}_1, \dots, \mathcal{P}_T)$ un multi-ensemble ordonné (chronologiquement) de T partitions du même espace D . Soit $\{R_i^t\}_{i=1}^{|\mathcal{P}_t|}$ l'ensemble des régions de la partition \mathcal{P}_t . Le SOTAG de $(\mathcal{P}_1, \dots, \mathcal{P}_T)$ est un graphe $\mathcal{X} = \mathcal{U} \times \mathcal{H}$ dont les nœuds $u \in \mathcal{U}$ sont les régions de la suite de partitions, et les arcs $h \in \mathcal{H}$ représentent la superposition de deux régions successives. On note \mathbf{R} l'application (bijective) qui à un nœud de \mathcal{X} associe une région de l'ensemble des partitions associé au graphe :

$$\begin{aligned} \mathbf{R} &: \mathcal{U} \longrightarrow \{\mathcal{P}_1, \dots, \mathcal{P}_T\} \\ u &\longmapsto R_i^t \end{aligned} \tag{4.3}$$

On peut aussi considérer une version attribuée de ce graphe dans laquelle les nœuds correspondent à des objets dynamiques de durée nulle. Ce sont des objets statiques avec un attribut supplémentaire indiquant l'instant d'existence. On appellera aussi ce graphe SOTAG. Les objets considérés étant caractérisés par un seul instant, les relations temporelles entre objets se résument aux relations *précède* et *égal* et à leurs spécialisations spatiales. La spécialisation spatiale considérée pour *précède* est la relation d'adjacence temporelle. L'égalité de deux objets correspond à l'égalité de leurs attributs, et on peut aussi considérer une relation de simple concomitance entre objets.

Notons que le SOTAG est un sous-graphe du graphe intersection. Par ailleurs, la relation "se superpose" n'étant pas transitive, l'existence d'un chemin $\text{path}(u_i, u_j)$ dans le SOTAG n'implique pas que $\mathbf{R}(u_i) \cap \mathbf{R}(u_j) \neq \emptyset$.

La figure 4.2 illustre le SOTAG.

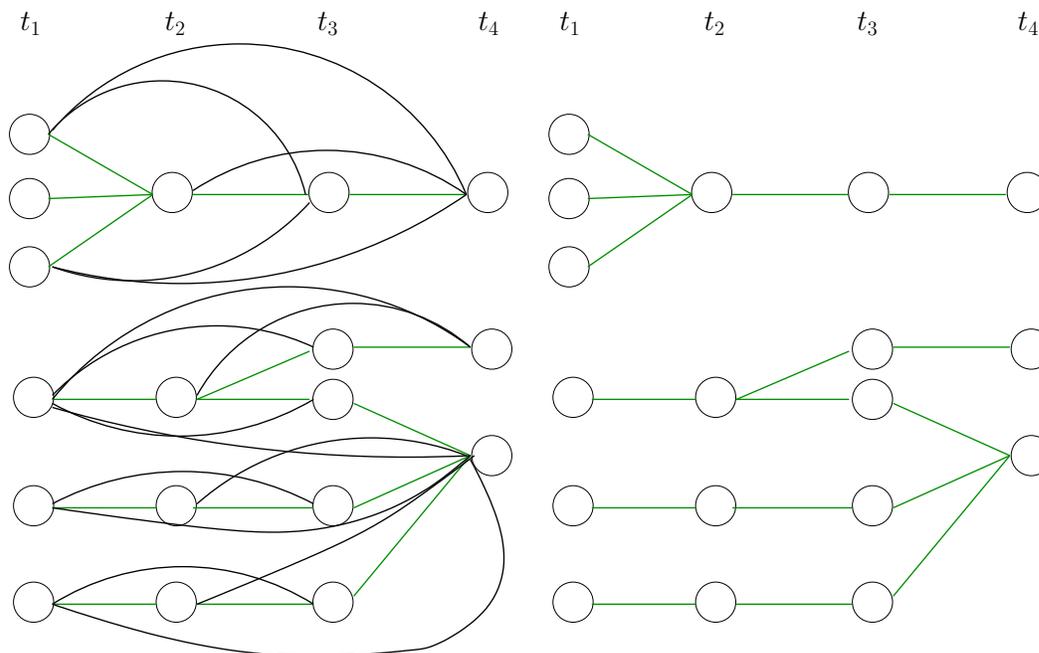


FIGURE 4.2 – Représentations du graphe intersection, à gauche, et de son sous-graphe, le graphe d’adjacence temporelle de régions (SOTAG), à droite.

4.4.2 Présentation de la méthode d’estimation du modèle de STIS

Le modèle d’STO que nous avons décrit suppose des STO de durée temporelle la durée de la STIS. Ces STO correspondent à des chemins dans le SOTAG.

Nous proposons dans un premier temps d’extraire le SOTAG. L’extraction du SOTAG consiste en l’extraction des régions spatiales définissant les différentes instances des STO et en la mise en correspondance de ces différentes instances. Nous décomposons donc cette partie en deux étapes : une étape de segmentation des images de la STIS en objets satisfaisant des contraintes sur la simplicité du SOTAG, et une fois les nœuds du SOTAG déterminés, une estimation des arcs du SOTAG.

A ce niveau, nous avons estimé les changements spatiaux des STO, mais leurs processus de caractéristiques restent inconnus. Nous proposons donc dans un second temps une analyse des caractéristiques des STO.

4.4.2.1 Segmentation jointe des STIS en instances d’OST

Nous avons vu que les changements topologiques découlent de la propriété d’homogénéité spatiale des caractéristiques des STO. Leur détermination image par image de façon indépendante par un algorithme de segmentation devrait donc suffire. Cependant, l’hypothèse de scène arrière-plan que nous avons faite n’est pas parfaitement vérifiée. En effet, cette hypothèse suppose que tout ce qui est visible dans la STIS provient de l’arrière-plan. Par exemple, les variations radiométriques de l’atmosphère s’ajoutent aux radiométries des objets dans la STIS, et le modèle, plutôt que de séparer les contributions, les attribue chacune à l’arrière-plan. Ainsi, une brume aura pour effet de diminuer fortement les contrastes de l’image, et donc des SO de notre modèle approximé. De plus, l’hypothèse d’homogénéité spatiale des caractéristiques des SO n’est pas exacte, et cet écart à la réalité devra être

pris en compte (par un bruit de caractéristiques ou par un seuillage). Des segmentations indépendantes ne fournissent donc pas les bons SO. Il nous faut donc exploiter une homogénéité temporelle supplémentaire. Par ailleurs, une homogénéité temporelle est nécessaire à la mise en correspondance des SO de façon à obtenir des STO. Nous avons vu au §A.2.3.3 différents types d’homogénéités : l’homogénéité de forme, de mouvement, de caractéristiques radiométriques, de profondeur, et de sémantique. Les homogénéités de mouvement et de profondeur ne présentent aucun intérêt pour les objets du fond, et l’homogénéité sémantique implique une intervention de l’utilisateur que nous souhaitons repousser le plus tard possible.

De plus, il n’y a pas d’homogénéité des caractéristiques car celles-ci sont supposées dynamiques pour les objets de l’arrière-plan. De plus, les discontinuités de caractéristiques ne peuvent être exploitées pour extraire les différentes instances d’STO car non seulement leurs événements ne correspondent pas nécessairement à des discontinuités, mais surtout, elles sont indépendantes des événements spatiaux. En effet, les événements spatiaux sont dus au fait que les caractéristiques changent et pas au fait qu’elles subissent des événements. Ces événements spatiaux auraient lieu de la même façon si les caractéristiques des STO n’étaient modifiées par aucun événement. Cette homogénéité n’est donc d’aucune utilité pour la détermination des instances des STO et des événements spatiaux qu’ils subissent.

En revanche, l’homogénéité de forme est applicable. En effet, dans notre modèle, les régions des objets de l’arrière-plan ne subissent aucun déplacement ou déformation. Cette homogénéité n’est cependant pas l’homogénéité traditionnelle qu’on suppose entre deux instances d’un objet dynamique en raison des changements topologiques des STO. L’homogénéité de forme considérée concerne des groupes d’objets bien choisis : en notant R_i^t la région spatiale associée à l’instance de l’STO O_i à l’instant t , une division de région se traduit par l’homogénéité de forme suivante : $R_i^t \simeq_S R_{i_1}^{t'} \cup \dots \cup R_{i_k}^{t'}$, réciproquement pour la fusion, et une combinaison de fusions et de divisions se traduit par l’homogénéité de forme suivante : $R_i^t \cup \dots \cup R_{i_k}^t \simeq_S R_{i_1}^{t'} \cup \dots \cup R_{i_l}^{t'}$, S étant la distance de forme considérée. En effet, en raison des bruits géométriques considérés, il n’y a pas égalité des formes. Ces points seront repris et détaillés dans le ch.9 dans lequel une méthode de segmentation jointe de la STIS est proposée.

Le tableau 1.1 compare notre méthode aux méthodes présentées dans l’état de l’art. Nous avons coché que notre modèle ne prend pas en compte des déformations d’objet. Ce point est ambigu dans la mesure où l’on ne cherche effectivement pas à estimer la déformation géométrique des objets. Cependant, une fusion ou une division contient implicitement une déformation de l’objet et afin d’être robuste aux erreurs de recalage, de faibles déformations des objets sont autorisées. Cette redondance de forme sera à nouveau exploitée dans le ch.8 dans lequel les SO précédemment déterminés sont mis en correspondance de façon à obtenir le SOTAG.

4.4.2.2 Analyse des caractéristiques des OST

Le graphe contient implicitement les événements spatiaux, et permet donc une analyse géométrique de la séquence. Les STO sont cependant aussi caractérisés par leurs caractéristiques (radiométriques) et leurs changements, et comme nous l’avons présenté dans le chapitre 3, l’analyse des évolutions radiométriques présente un intérêt particulier pour les STIS. Nous proposons donc dans le ch.9 une analyse radiométrique des STO.

En raison des évolutions radiométriques locales des objets de la scène, les images ne présentent pas d’homogénéité radiométrique globale et en conséquence, les approches clas-

siques d'inter-calibration ne sont pas applicables. De plus, la fréquence d'acquisition est faible par rapport aux phénomènes observés et des superpositions par des nuages, non prises en compte par le modèle sous l'hypothèse d'arrière-plan insèrent dans les signaux temporels des valeurs aberrantes. Le nombre de ces valeurs aberrantes est renforcé par l'approche séquentielle adoptée : extraction des objets spatiaux, suivie de leur mise en correspondance, et finalement la caractérisation de leur évolution. Enfin, la scène est hétérogène, et la nature des objets n'est pas connue *a priori*.

En conséquence, les approches par filtrage s'appuyant sur des hypothèses de continuité ou utilisant des *a priori* ne sont pas applicables, et les approches par détection de changements ne permettent pas de caractériser les évolutions elles-mêmes. Nous proposons donc de compenser le manque d'information par les redondances intra-classes. Ce point de vue est similaire au point de vue adopté lors de l'analyse de STIS-CR (§A.2.4.1) dont les différents types d'objets sont classifiés selon des signatures temporelles. Cependant, afin de prendre en compte des successions différentes de processus pour les différents STO, et afin d'être robuste aux divergences locales des groupements d'une classe sémantique liées aux valeurs aberrantes, nous supposerons un support temporel non fixe et variable. L'extraction de ces motifs fréquents permet de capturer automatiquement des regroupements statistiquement pertinents.

4.5 Conclusion

Nous avons dans ce chapitre proposé un modèle de scène de télédétection à haute résolution spatiale et temporelle. Ce modèle permet une description bas-niveau de la STIS et différentes interprétations liées à différents contextes sémantiques peuvent en découler. Il s'agit donc d'une modélisation de STIS-HR comparable à la segmentation d'image dans sa finalité et dans son utilisation qui est indépendante de l'application.

Ce modèle est constitué d'objets spatio-temporels de différents plans. Un arrière-plan dont les objets évoluent radiométriquement et dont les régions sont en conséquence modifiées par des événements spatiaux de type fusion et division, et des premiers plans constitués d'objets invariants radiométriquement, mais dont les régions changent via des processus spatiaux de type mouvement rigide et déformation. En raison de la complexité de l'estimation du modèle, et du sous-échantillonnage temporel des STIS actuelles par rapport à certains phénomènes, nous avons ensuite supposé que la scène observée ne contenait que l'arrière-plan. Afin de décrire toutefois la géométrie de la scène de façon complète, nous considérons dans le modèle approximé de faibles déformations et déplacements des régions qui seront prises en compte dans un bruit géométrique, et des déplacements ou déformations importantes décrites via un événement spatial particulier : la combinaison de fusions et de divisions. L'arrière-plan est modélisé par un graphe d'adjacence temporel des instances des objets spatio-temporels de sorte que ces derniers correspondent à des chemins dans le graphe.

Chapitre 5

Segmentation

5.1 Différentes approches pour la segmentation d'images

Pour des séquences d'images comme pour des images, l'objectif de la segmentation est d'extraire les objets présents dans les données. Cette approche est donc basée sur l'hypothèse que les objets sont des régions (ou des groupes de régions,¹) dont une certaine caractéristique est homogène. Cette formulation considère implicitement un modèle d'image. Un certain nombre d'approches considèrent ce modèle de façon implicite en exploitant des heuristiques afin de mettre en évidence les régions, et d'autres considèrent le modèle de façon explicite. La première classe de méthode est généralement locale alors que la seconde est globale (du moins dans la formulation du problème). Deux points de vue duaux peuvent ensuite être considérés. Le point de vue *contour* qui exploite d'avantage la disparité inter-objets, et le point de vue *région* qui exploite davantage l'homogénéité des régions.

Dans les approches adoptant le point de vue contour avec le modèle implicite, il y a toutes les approches par filtrage utilisant des opérateurs différentiels. Nous citerons en particulier le filtre de Canny [Canny, 1986]. Les contours obtenus par filtrage peuvent cependant être épais, et des filtres morphologiques d'amincissement sont classiquement utilisés en post-traitement. Le but de la segmentation est cependant d'extraire des objets, or la forme des objets sera très sensible à des erreurs locales de détection. Afin de réduire ce défaut, des filtres morphologiques de fermeture sont généralement appliqués. Une autre approche permettant de gérer les problèmes de fermeture des régions et d'amincissement des contours est la méthode de segmentation par ligne de partage des eaux (LPE) [Beucher and Lantuejoul, 1979, Vincent and Soille, 1991]. Dans ce cas, l'image est considérée comme un paysage qu'on inonde. Les *bassins* représentent les régions et sont séparés par des digues qui représentent les contours. Ce résultat est généralement intermédiaire et différentes utilisations sont possibles. Ainsi, dans [Beucher and Lantuejoul, 1979, Meyer, 2001], des marqueurs sont utilisés pour indiquer les zones d'intérêt, et la LPE permet de localiser les frontières englobant le marqueur de façon précise. D'autres approches considèrent un paramètre permettant de régler le niveau d'inondation du paysage. Ces approches permettent de construire une hiérarchie de segmentation qui peut être représentée par un arbre de fusion des régions dans lequel des coupes sont effectuées pour obtenir les segmentations d'intérêt [Y.Lemur, 2004, Gueguen, Le Men and Datcu, 2006, Giros, 2006]. Le lien entre la segmentation par LPE et la formulation variationnelle a été mis en évi-

1. Dans ce cas, un post-traitement de regroupement doit être effectué.

dence dans [Nguyen et al., 2003] en considérant une distance topographique, et un terme de régularisation supplémentaire est considéré pour construire des *watersnakes*.

Dans les approches considérant toujours implicitement le modèle, mais adoptant un point de vue région, nous trouvons les méthodes de croissance de région, de fusion, de division et de fusion et division. Dans ce cas, la segmentation est vue comme un problème de partitionnement des images en régions vérifiant des prédicats d'homogénéité [Chang and Li, 1994, Horaud and Monga, 1995, Horowitz and Pavlidis, 1976]. Il existe différents prédicats et les plus fréquemment utilisés sont présentés dans [Maitre, 2003]. Les critères d'homogénéités considérés correspondent à des modèles probabilistes particuliers et des liens avec la théorie bayésienne et le MDL ont été mis en évidence dans [Zhu et al., 1995].

Ces approches s'appuient sur des heuristiques issues d'un modèle implicite de l'image. Une autre classe d'approches considère une modélisation explicite de l'image de façon à minimiser une énergie globale. Parmi ce type de méthodes, on distingue les approches variationnelles, les approches bayésiennes et les approches par codage minimal qui sont celles auxquelles on s'intéressera tout particulièrement. Pour ces approches par minimisation d'énergie, il existe les mêmes deux points de vues région et contour.

Les approches variationnelles sont issues de la formulation du problème de segmentation par [Mumford and Shah, 1989] dans lequel un modèle d'images lisses par morceaux est considéré. Ce point de vue a donné naissance à la segmentation par contours actifs [Kass et al., 1988] et par ensembles de niveaux [Osher and Sethian, 1988, Vese and Chan, 2002]. L'approche par contour actif considère un contour soumis à des forces liées à la valeur du gradient de l'image au niveau du contour et à une élasticité. Une extension de cette approche permettant les changements topologiques considère ces contours comme des coupes d'une surface dans un espace tridimensionnel, et l'approche par ensembles de niveaux permet une optimisation rapide de cette surface vers son état d'équilibre. L'énergie considérée n'est donc pas locale, en revanche, l'optimisation l'est généralement, et de tels snakes peuvent donc délimiter des régions relativement hétérogènes et sont très sensibles à l'initialisation. Une extension de cette approche a été proposée dans [Paragios and Deriche, 2002] de façon à prendre justement en compte l'intérieur des régions et assurer ainsi leur homogénéité. Une telle formulation se rapproche donc plus des approches région. Dans ce cas, l'écart aux données est considéré de façon déterministe, en calculant une distance (généralement L_2) entre les données et les données ajustées.

Dans le cas des approches bayésiennes, le problème de la segmentation est vu comme un problème inverse de restauration d'image qui peut se résoudre par maximisation a posteriori des données sachant le modèle [Peng and Pieczynski, 1995, Comaniciu and Meer, 1997]. Différentes optimisations sont alors possibles et en particulier, le formalisme des champs de Markov est particulièrement adapté car il permet de prendre en compte des contraintes spatiales et des optimisations efficaces à ces problèmes existent [Geman and Geman, 1984, Hochbaum, 2001].

Un point critique de la segmentation d'images est de trouver le juste compromis entre la régularisation et la fidélité. Ce compromis, qui est lié au choix de l'échelle spatiale, est souvent subjectif et dépend des applications. Aussi, une variété d'approches proposant des hiérarchies de segmentations ont été proposées. Certains algorithmes contiennent implicitement cette hiérarchie par la manière dont sont obtenues les fusions, comme la segmentation par LPE mentionnée ci-dessus. Cette idée de hiérarchie de segmentation est reprise dans de nombreux algorithmes [Koepfler et al., 1994, Guigues, 2003, Haxhimusa and Kropatsch, 2004]. Enfin d'autres approches consistent à rechercher les coupes minimales dans le graphe d'adjacence des pixels aux arcs pondérés de l'image [Boykov and

Funka-Lea, 2006, Wu and Leahy, 1993, Shi and Malik, 2000]. Enfin, il existe l'approche de coupes normalisées proposée dans [Shi and Malik, 2000].

Nous détaillons maintenant les approches par codage minimal qui sont celles que nous avons adoptées par la suite.

5.2 Segmentation d'images par minimisation de la longueur de description

La segmentation d'une image en régions peut être vue comme un problème de modélisation des données, le modèle considéré étant un ensemble de sous-images homogènes dont les supports spatiaux (régions) forment une partition de l'espace. Il n'y a pas unicité d'un tel modèle, cependant, généralement, la partition d'intérêt est celle qui permet un juste compromis entre simplicité du modèle, et fidélité du modèle aux données. Le principe de minimisation de la longueur de description (MDL) fut introduit par Rissanen en 78. Ce principe est rappelé dans l'annexe A.1. Il s'agit d'un principe inspiré de la théorie du codage. Le problème est reformulé de la façon suivante : on souhaite transmettre sans perte les données de la façon la moins coûteuse. Cette transmission nécessite deux messages : une description du modèle des données, et les données décrites à l'aide du modèle. La longueur de ces deux messages (en bits) mesure de façon homogène deux aspects complémentaires des données : le contenu déterministe (porté par le modèle) et la vraisemblance de ce modèle. Plus le modèle est complexe, plus il permettra de décrire fidèlement les données, permettant ainsi une deuxième partie du message courte. Cependant, la complexité du modèle entraînera une première partie de message très longue. La minimisation de la longueur totale permet donc d'obtenir un compromis entre ces deux contraintes.

Cette approche permet donc de sélectionner parmi les différentes modélisations possibles la plus simple qui est aussi fidèle aux données. Comparée à d'autres approches (BIC, Akaike) qui utilisent une pénalisation par la longueur de description, elle présente l'intérêt de mettre automatiquement à l'échelle différents types de modèles. Ainsi, des objets de la scène, qui sont caractérisés par une région spatiale, et un vecteur de caractéristiques nécessiteront au moins deux modèles différents. Enfin, la mise à l'échelle des mesures des différentes parties du modèle rend la méthode de segmentation automatique.

Le MDL a été appliqué au problème de la segmentation d'images pour la première fois dans [Leclerc, 1989]. Le problème de la segmentation y est présenté comme un problème inverse, en considérant que l'image observée, Im est le résultat d'un bruitage sur une image constante par morceau ou lisse par morceaux, $\tilde{\text{Im}}$:

$$\text{Im} = \tilde{\text{Im}} + n \quad (5.1)$$

où n représente un bruit additif. Ce problème est ensuite traité comme un problème de codage au cours duquel les différentes parties du messages sont codées en différents sous-messages : $\mathcal{L}(\text{Im}) = \mathcal{L}_i(\tilde{\text{Im}}) + \mathcal{L}_n(n)$. \mathcal{L}_i , et \mathcal{L}_n correspondent aux longueurs de description dans des *langues* adaptées, respectivement à l'image d'origine, et au bruit. Ces langues représentent la manière de décrire, elles fournissent un vocabulaire. Prenons l'exemple d'une image constante par morceaux. La restriction de l'image à une région est appelée sous-image. La langue permettant de traduire simplement cette image consiste à décrire, pour chaque sous-image, ses contours, puis sa valeur. Choisir ces langues revient donc à choisir un a priori. La stabilité de la solution provient du codage en plusieurs parties. Ainsi, la décomposition du message associé à l'image en plusieurs messages transportant séparément

l'information concernant la partition, les valeurs moyennes ou paramètres radiométriques, et le bruit permet que la modification d'un de ces messages n'altère pas les autres.

De façon générale, le modèle de la segmentation consiste en une image sous-jacente homogène par morceaux. Il peut donc être décrit par une paire $(\mathcal{P}, \mathbf{J})$, où \mathcal{P} est une partition du domaine de l'image, D (c'est-à-dire qu'elle est constituée d'un ensemble d'ouverts $\{R_i\}_{i=1}^{|\mathcal{R}|}$, et de sa fermeture \mathcal{C} dans D) et où $\mathbf{J} = \{J_i\}$ est l'ensemble des modèles de sous-images associés à chacune des régions. :

$$\tilde{Im} = (\mathcal{P}, \mathbf{J}) \quad (5.2)$$

La longueur de description totale s'écrit donc $\mathcal{L} = \mathcal{L}_{\text{geom}}(\mathcal{P}) + \mathcal{L}_{\text{radiom}}(\mathbf{J}) + \mathcal{L}_{\text{noise}}(\tilde{Im}|\text{Im})$. Le bruit et les valeurs déterministes des sous-images sont généralement regroupées en des sous-images dont les valeurs sont les réalisations de processus stochastiques de moyenne non nulle. La description de chaque sous-image nécessite donc la description des paramètres de la distribution considérée, ainsi que la longueur de description des valeurs selon ce modèle. En notant Ψ les paramètres des sous-images, on a donc :

$$\mathcal{L} = \mathcal{L}_{\text{geom}}(\mathcal{P}) + \mathcal{L}_{\text{param}}(\Psi) + \mathcal{L}_{\text{fid}}(\text{Im}|\mathcal{P}, \Psi) \quad (5.3)$$

Dans [Nicolas et al., 2001], une approximation est aussi considérée pour le codage des contours. On considère ici que ces contours sont exacts et codés de façon déterministe. Les erreurs sur les contours sont alors directement répercutées sur la longueur de description de fidélité des valeurs de la région.

5.2.1 Longueur de description

Détaillons donc les différentes longueurs de description géométrique, $\mathcal{L}_{\text{geom}}(\mathcal{P})$, des paramètres, $\mathcal{L}_{\text{param}}(\Psi)$, et de fidélité $\mathcal{L}_{\text{fid}}(\text{Im}|\mathcal{P}, \Psi)$ proposées dans la littérature.

5.2.1.1 Longueur de description de la région spatiale

Dans [Leclerc, 1989], les contours de la partition sont décrits à l'aide d'un code de chaîne des contours : pour chaque région, on doit coder la position d'un point de départ, puis, la suite des directions à prendre pour parcourir totalement le contour. Cette longueur, sommée sur toutes les régions est divisée par deux car chaque contour est décrit une fois au titre de sa région extérieure et une autre fois au titre de toutes ses régions environantes.

En considérant que les contours sont des chemins orientés sur la grille duale de la grille des pixels, le codage du premier nœud revient à choisir une position parmi les $N_x * N_y$ positions de l'image. La longueur de description correspondante est donc $\log(N_x * N_y)$. En considérant une 4-connexité, l'arc menant au prochain nœud du contour est l'un des quatre arcs sortant du nœud initial. La longueur de description associée est donc $\log(4)$. Puis, les $|\mathcal{C}| - 2$ prochains nœuds du contour \mathcal{C} sont atteints en choisissant l'un des arcs sortants du nœud courant sans revenir en arrière. La longueur de description de chacune de ces directions est donc $\log(3)$. La longueur de description résultante pour une région est donc la suivante :

$$\mathcal{L}_{\text{geom}}(R_i) = \frac{1}{2}(\log(N_x * N_y) + \log(4) + (|\mathcal{C}_i| - 2) * \log(3)) \quad (5.4)$$

Dans [Kanungo et al., 1994, Nicolas et al., 2001, Galland et al., 2003], les contours sont représentés à l'aide d'un graphe dont les arcs représentent des portions de contours

séparant deux régions, et les nœuds les jonctions entre ces portions de contours, ie : les contours communs à plus de deux régions. Le parcours complet du graphe nécessite le parcours de chacune de ses composantes connexes. Pour chaque composante connexe, il existe un parcours permettant de visiter une seule fois chaque arc si et seulement si la composante connexe considérée est un Graphe eulérien (voire A.7). Cependant, dans le cas général, tel n'est pas le cas, et on décompose alors la composante connexe en sous-graphes eulériens. Le parcours de la composante connexe nécessite donc autant de points de départ que de graphes eulériens qui la composent. Le codage des contours consiste donc à coder la position du pixel associé à chaque nœud de départ, puis à coder tous les arcs. Dans [Kanungo et al., 1994], le codage des points de départ est négligé. [Galland et al., 2003] propose une expression exacte de la longueur d'une telle description en remarquant que le nombre de points de départ est égal au nombre de composantes connexes sans nœud de degré impair plus la moitié du nombre de nœuds de degré impair. En effet, un nœud de degré impair est soit un nœud de départ, soit un nœud d'arrivée. [Nicolas et al., 2001] utilise une description encore plus compacte en utilisant des clés dans un code de chaîne du graphe précédant chaque nœud de départ. Le parcours de cette chaîne en stockant chaque clé dans une file FIFO (First In, First Out) permet de revenir au point de départ d'un chemin que l'on vient de finir de parcourir afin de terminer le parcours du précédent chemin là où on l'avait laissé. Ce codage remplace donc le codage de la position de chaque point de départ par le codage d'une clé. Pour le codage des arcs du graphe, [Kanungo et al., 1994] considère une description exacte du contour, et utilise la même description par codage de chaîne que dans [Leclerc, 1989]. La longueur de chaque arc doit donc être codée, et le codage de prior universel des entiers [Rissanen, 1983] est utilisé pour mesurer la longueur de description associée. [Galland et al., 2003] et [Nicolas et al., 2001] considèrent des contours polygonaux, et chaque arc est une ligne droite. L'offset en x et en y doit donc être codé pour chaque arc. [Galland et al., 2003] considère une distribution normale de ces offsets dont il code aussi les paramètres de façon à utiliser la longueur de description optimale de Shannon.

Les formes contenant des portions rectilignes couteront donc moins cher dans le système de codage de [Galland et al., 2003] que dans celui de [Leclerc, 1989] ou [Kanungo et al., 1994]. D'autre part, le codage de [Galland et al., 2003] est moins sensible au changement de résolution. Ainsi, une région carrée de 1 pixel dans une image de taille $N_x * N_y$ coûte $\log(N_x * N_y) + \log(4) + 2\log(3)$, et $\log(s^2 N_x * N_y) + \log(4) + (4s - 2)\log(3)$ pour une image s fois mieux résolue alors que la longueur de description de [Galland et al., 2003] reste identique. Et la longueur de description de [Kanungo et al., 1994] est légèrement plus faible que celle considérée par [Leclerc, 1989].

5.2.1.2 Longueur de description des paramètres

La partie géométrique de \tilde{I}_m ainsi décrite, les valeurs des sous-images doivent être codées. Deux a priori sont considérés dans [Leclerc, 1989] : une image constante par morceaux, et une image lisse par morceaux. Dans le cas de sous-images constantes, un seul paramètre par région doit être décrit. Ce paramètre, qui est la moyenne, μ_i des pixels sur chaque région a donc pour longueur [Rissanen, 1989] : $\mathcal{L}(\mu_i) = \frac{d}{2} \log(R_i)$, en notant d , la dimension. Dans le cas d'un bruit gaussien de moyenne nulle uniforme par morceaux, comme considéré dans [Leclerc, 1989, Kanungo et al., 1994], on a $\Psi_i = (\mu_i, \Gamma_i)$, où Γ_i est la matrice de covariance de la distribution gaussienne de la région. La longueur de description

des paramètres s'écrit donc dans ce cas :

$$\mathcal{L}_{\text{param}}(\Psi_i) = d(1 + \frac{d+1}{2}) \log(|R_i|) \quad (5.5)$$

Cette modélisation du bruit rend partiellement compte des différentes textures.

Dans le cas de sous-images à valeurs sur une surface polynomiale [Leclerc, 1989, Kanungo et al., 1994], les sous-images ont une moyenne qui varie spatialement. Les coefficients non nuls de la moyenne polynomiale, ainsi que la matrice de covariance doivent donc être codés. Dans le cas multidimensionnel, la longueur de description de ces coefficients est donc : $\mathcal{L}(\alpha_i) = \frac{d}{4} \cdot (p_i + 1)(p_i + 2) \cdot \log(|R_i|)$ où p_i est le degré du polynôme de la moyenne de la région R_i . En toute rigueur, il faudrait aussi coder ces degrés, cependant, [Kanungo et al., 1994] montre expérimentalement que la longueur associée peut être négligée. Donc, pour tous les paramètres, on obtient : $\mathcal{L}_{\text{param}} = \frac{d}{4} ((d+1) + (p_i+1)(p_i+2)) \log(|R_i|)$.

Dans [Galland et al., 2003] le problème est étendu au cas de bruits gamma afin de rendre compte du speckle des images radar, puis dans [Galland et al., 2005] des bruits de Poisson et Bernoulli sont aussi considérés. Le codage des paramètres se fait exactement comme précédemment, en notant k , le nombre de paramètres, on a $\mathcal{L}_{\text{param}} = \frac{k}{2} \log(|R_i|)$. [Delyon et al., 2006] présente une extension au cas où les pixels des régions suivent des distributions de type inconnu. Les pixels de chaque région sont vus, comme précédemment comme des réalisations de variables aléatoires indépendantes et identiquement distribuées selon une distribution P_r qui est ici inconnue, et ne fait pas nécessairement partie des distributions canoniques. La distribution est alors modélisée par une fonction en escalier de Q marches de longueurs variables qui approximent l'histogramme de la région. Les paramètres à coder sont dans ce cas la largeur des bins de l'histogramme de la région, ainsi que chacune des probabilités qui leur sont associées. Cet algorithme permet donc de segmenter automatiquement des images dont les régions peuvent avoir des textures très différentes.

5.2.1.3 Terme de fidélité correspondant à une région spatiale

La description du bruit découle directement de la longueur de code optimale de Shannon : $\mathcal{L}(x) = -\log(P(x))$. Donc $\mathcal{L}_{\text{fid}}(R_i) = -\log(P(\mathbf{Im}_{R_i} | \Psi_i))$. Nous considérons que les pixels d'une région R_i sont des réalisations indépendantes d'une même variable aléatoire de paramètre Ψ_i . Dans le cas d'un modèle de sous-images de moyennes μ_i et de matrices de covariance Γ_i , la probabilité de la restriction de l'image, \mathbf{Im}_{R_i} à la région spatiale R_i est donc :

$$P(\mathbf{Im}_{R_i}; \mathcal{N}(\mu_{R_i}, \Gamma_i)) = \prod_{j=1}^{|\mathbf{Im}_{R_i}|} \frac{1}{(2\pi)^{\frac{3}{2}} \sqrt{|\Gamma_{R_i}|}} \exp\left(-\frac{(\mathbf{Im}_{R_i}(j) - \mu_i)^T \Gamma_i^{-1} (\mathbf{Im}_{R_i}(j) - \mu_i)}{2}\right) \quad (5.6)$$

En remarquant que $\mu_i = \frac{1}{|\mathbf{Im}_{R_i}|} \sum_{j=1}^{|\mathbf{Im}_{R_i}|} \mathbf{Im}_{R_i}(j)$, et $\Gamma_{R_i} = \frac{1}{|\mathbf{Im}_{R_i}|} (\mathbf{Im}_{R_i}^T \mathbf{Im}_{R_i}) - \mu_i^T \mu_i$, la longueur de description du terme de fidélité radiométrique devient :

$$\mathcal{L}_{\text{fid}}(R_i) = \frac{|\mathbf{Im}_{R_i}|}{2} (3 + \log(2\pi) + 3 \cdot \log(|\Gamma_{R_i}|)) \quad (5.7)$$

5.2.2 Optimisation

La longueur de description complète de l'image est ainsi déterminée. Dans [Leclerc, 1989], afin d'optimiser la longueur de description totale, on la plonge dans un espace

continu en remplaçant les distributions de Dirac modélisant les changements d'une région à l'autre par des exponentielles dont on réduit progressivement le facteur d'échelle jusqu'à convergence. A chaque itération (pour chaque facteur d'échelle), on retrouve un système linéaire que l'on résout à l'aide de la méthode de Seidel. On obtient ainsi un minimum global, en revanche, dans le cas de non convergence, des régions peuvent être non fermées.

Dans le cas multidimensionnel, l'espace de recherche est trop grand pour envisager une telle méthode. [Kanungo et al., 1994] a donc proposé une méthode d'optimisation par fusion de régions afin de traiter le cas multidimensionnel. On utilise pour cela une pile prioritaire permettant d'ordonner les paires de régions candidates à fusion par coût de fusion croissant. A chaque étape, la paire de régions correspondant au coût le plus négatif est fusionnée, et ceci tant que le coût est négatif, ce qui correspond à une décroissance de la longueur de description. Cette méthode permet d'obtenir un minimum local.

Dans [Galland et al., 2003], l'utilisation d'une grille polygonale que l'on modifie en enlevant des arcs et en le déformant permet une optimisation globale. La grille considérée est un ensemble de nœuds séparés par des arcs permettant de délimiter les régions. Afin de délimiter les régions sans ambiguïté, cette grille est décalée de $(1/2, 1/4)$ par rapport à la grille des pixels. D'abord la grille est initialisée (par exemple par la grille des pixels décalée), puis une séquence de trois opérations est effectuée : la fusion de régions, le déplacement de nœuds, et la suppression de nœuds. L'étape de fusion de nœud est similaire à celle de [Kanungo et al., 1994]. Ensuite, un déplacement aléatoire de chaque nœud de la grille est testé, et il est accepté si il correspond à une décroissance de la longueur de description. Finalement, la suppression de chaque nœud de la grille de degré 2 est testée et acceptée si elle résulte en une décroissance de la longueur de description. Cette dernière étape permet une régularisation du contour. Notons qu'à chacune de ces étapes, le calcul de la longueur de description des données conditionnées par le modèle nécessite le calcul de ses statistiques suffisantes qui requiert une sommation sur les régions. Afin de réduire sensiblement le temps de calcul, [Galland et al., 2003] propose un pré-traitement permettant de remplacer la sommation sur la région par une sommation sur les contours de cette image prétraitée.

5.2.2.1 Détail de l'implémentation d'une segmentation par MDL optimisée par fusion

Pour des raisons que nous éclaircirons au §7.5, ce type d'optimisation ne peut être étendu à la longueur de description que nous allons proposer pour la STIS. Nous utiliserons donc une méthode d'optimisation par fusions, similaire à celle proposée dans [Kanungo et al., 1994]. Afin de mettre en évidence les similitudes et les différences avec le cas 2D, nous proposons de détailler ici l'implémentation de la méthode de [Kanungo et al., 1994] pour un modèle d'image décomposé en sous-images dont les valeurs suivent une distribution gaussienne. D'après les équations 5.4, 5.5, et 5.7, la longueur de description correspondante s'écrit :

$$\mathcal{L} = \frac{|P|}{2} (\log(N_x * N_y) + \log(4)) + N_x * N_y (3 + \log(2\pi)) + \sum_{i=1}^{|P|} \left((|C_i| - 2) * \log(3) + \frac{9}{2} \log(|R_i|) + \frac{|\mathbf{Im}_{R_i}|}{2} (\log(|\Gamma_{R_i}|)) \right) \quad (5.8)$$

On initialise d'abord la partition avec une sur-partition, puis, pour chaque paire de région voisine, on calcule la décroissance de longueur de description, $\delta\mathcal{L}$. Tant qu'il existe une fusion permettant une réduction de la longueur de description, on la fusionne et on met à jour les coûts de fusion correspondants.

Etant donné que seules les différences de longueurs de descriptions nous intéressent, nous proposons de calculer directement ces différences sans passer explicitement par les

longueurs de descriptions.

Une décroissance doit être calculée pour toutes les paires de régions spatialement adjacentes. On utilise donc la structure de graphe d'adjacence de régions spatiales, \mathcal{SG} dont les arcs représentent toutes les paires de régions adjacentes. Une version attribuée de ce graphe, notée \mathcal{ASG} , contient donc toutes les données nécessaires pour le calcul des décroissances liées à la fusion. Les attributs considérés pour les nœuds sont les statistiques suffisantes d'ordre 1 et 2 et la taille (nombre de pixels) de la région correspondante, et les attributs pour les arcs sont le nombre de segments de la portion de contour séparant les régions associées aux deux extrémités de l'arc.

Calcul des variations de longueur de description D'après l'équation 5.3, la décroissance totale à considérer est :

$$\delta_{it}^{it+1} \mathcal{L} = \delta_{it}^{it+1} \mathcal{L}_{\text{geom}} + \delta_{it}^{it+1} \mathcal{L}_{\text{param}} + \delta_{it}^{it+1} \mathcal{L}_{\text{fid}} \quad (5.9)$$

Détaillons maintenant les différentes composantes géométriques, paramétriques, et de fidélité radiométrique du gain de fusion.

Considérons la partition \mathcal{P}^{it} correspondant à la partition de l'image à l'itération it . A l'itération suivante, deux régions voisines R_i^{it} et R_j^{it} de cette partition ont été fusionnées de sorte que la partition de l'image à l'itération $it + 1$ est $\mathcal{P}^{it+1} = \mathcal{P}^{it} \setminus \{R_i^{it}, R_j^{it}\} \cup R_i^{it+1}$ où $R_i^{it+1} = R_i^{it} \cup R_j^{it}$. La décroissance de la longueur de description géométrique liée à la fusion correspondante est donc :

$$\delta_{it}^{it+1} \mathcal{L}_{\text{geom}} = \mathcal{L}_{\text{geom}}(\mathcal{P}^{it+1}) - \mathcal{L}_{\text{geom}}(\mathcal{P}^{it}) \quad (5.10)$$

Cependant, seules les régions R_i^{it} et R_j^{it} ont été modifiées donc :

$$\begin{aligned} \delta_{it}^{it+1} \mathcal{L}_{\text{geom}} &= \sum_{\substack{k=1 \\ k \neq i, j}}^{|\mathcal{P}^{it}|} \mathcal{L}_{\text{geom}}(R_k^{it}) + \mathcal{L}_{\text{geom}}(R_i^{it+1}) - \\ &\quad \left(\sum_{\substack{k=1 \\ k \neq i, j}}^{|\mathcal{P}^{it}|} \mathcal{L}_{\text{geom}}(R_k^{it}) + \mathcal{L}_{\text{geom}}(R_i^{it}) + \mathcal{L}_{\text{geom}}(R_j^{it}) \right) \\ &= \mathcal{L}_{\text{geom}}(R_i^{it+1}) - \mathcal{L}_{\text{geom}}(R_i^{it}) - \mathcal{L}_{\text{geom}}(R_j^{it}) \end{aligned} \quad (5.11)$$

Notons $\mathcal{C}(R_i^{it}, R_j^{it})$ représente la portion de contour séparant les régions R_i^{it} et R_j^{it} . Le contour de la réunion est égal à la réunion des contours privé de deux fois cette portion de contour commune. Donc d'après 5.4,

$\mathcal{L}_{\text{geom}}(R_i^{it+1}) = \frac{1}{2} \left(\log(N_x * N_y) + \log(4) + (|\mathcal{C}_i^{it}| + |\mathcal{C}_j^{it}| - 2|\mathcal{C}(R_i^{it}, R_j^{it})| - 2) \log(3) \right)$. La décroissance géométrique est donc :

$$\delta_{it}^{it+1} \mathcal{L}_{\text{geom}} = -\frac{1}{2} \left(\log(N_x * N_y) + \log(4) + (2|\mathcal{C}(R_i^{it}, R_j^{it})| + 1) \log(3) \right) \quad (5.12)$$

Le calcul de cette décroissance nécessite donc simplement le nombre de segments des contours communs. C'est pourquoi nous l'avons choisi comme attribut des arcs du graphe.

Lors d'une fusion, les paramètres des régions autres que la région fusionnée restent inchangés par rapport aux paramètres des régions avant fusion autres que les deux candidates à fusion. En conséquence, la décroissance de la longueur de description des paramètres consiste en la différence entre la longueur de codage des paramètres des deux régions candidates à fusion, et la longueur de codage des paramètres de la nouvelle région fusionnée. La décroissance correspondante est donc : $\delta_{it}^{it+1} \mathcal{L}_{\text{param}}(u_i, u_j) = \frac{|\Psi_i|}{2} \cdot \log\left(\frac{|R_i| + |R_j|}{|R_i| |R_j|}\right)$,

et en dimension 3, le nombre de paramètres des distributions radiométriques est 9 (3 valeurs moyennes et 6 valeurs pour la matrice de covariance), donc $\delta_{it}^{it+1} \mathcal{L}_{\text{param}}(u_i, u_j) = \frac{9}{2} \cdot \log\left(\frac{|R_i|+|R_j|}{|R_i||R_j|}\right)$.

Le calcul de cette décroissance nécessite donc simplement la taille des régions à fusionner. C'est pourquoi nous l'avons considéré comme attribut des nœuds du graphe d'adjacence de régions.

Dans le cas d'un modèle radiométrique gaussien, les statistiques suffisantes sont $G(u_i) = \sum_{k=1}^{|R_i|} \mathbf{x}_k^T \mathbf{x}_k (= \mathbf{X}_i^T \mathbf{X}_i)$, et $m(u_i) = \sum_{k=1}^{|R_i|} \mathbf{x}_k$, où $\mathbf{X}_i^T = [\mathbf{x}_1 \dots \mathbf{x}_{|R_i|}]$ représente l'ensemble des valeurs de la région R_i , et $\mathbf{x}_k^T = [x_{k1} \dots x_{kd}]$ représente la valeur (vectorielle) en un pixel. De même que pour la décroissance géométrique et paramétrique, le terme de fidélité n'est modifié que pour la région fusionnée, et consiste en la différence entre la longueur de la description de fidélité des régions candidates à fusion et la longueur de description de fidélité de la région fusionnée résultante. Contrairement aux autres variations, celle-là sera toujours positive. Et d'après l'équation 5.7 : $\delta \mathcal{L}_{\text{fid}}(u_i, u_j) = \frac{3}{2} \cdot (|R_{ij}| \log(|\Gamma_{ij}|) - |R_i| \log(|\Gamma_i|) - |R_j| \log(|\Gamma_j|))$. Or $\Gamma_i = \frac{1}{|R_i|} (G(u_i) - m(u_i)^T m(u_i))$. De plus, $G(u_{ij}) = G(u_i) + G(u_j)$, $m(u_{ij}) = m(u_i) + m(u_j)$, et $|R_{ij}| = |R_i| + |R_j|$. Donc :

$$\begin{aligned} \delta \mathcal{L}_{\text{fid}}(i, j) = \frac{d}{2} \cdot \left((|R_i| + |R_j|) \cdot \log\left(\frac{1}{|R_i|+|R_j|} \cdot (G(i) + G(j) + (m(i) + m(j))^T (m(i) + m(j)))\right) \right. \\ \left. - |R_i| \log\left(\frac{1}{|R_i|} (G(i) - m(i)^T m(i))\right) - |R_j| \log\left(\frac{1}{|R_j|} (G(j) - m(j)^T m(j))\right) \right) \end{aligned} \quad (5.13)$$

La croissance associée à la longueur de description de fidélité radiométrique se calcule donc à partir des statistiques suffisantes associées à chacune des régions candidates à fusion.

Chapitre 6

Recherche d'ensembles de symboles fréquents

La recherche de motifs fréquents est un domaine important de la fouille de données et de la découverte de connaissance dans les bases de données. Son point de départ est lié à l'analyse de *paniers d'articles* et spécialement à la fouille de base de transactions dans le but de décrire le comportement des clients de grands magasins. Un nombre important d'algorithmes a donc été proposé pour répondre à ce problème généralement connu sous le nom de fouille d'ensemble d'articles,¹ dont les plus connus sont Apriori, ECLAT, et FP-Growth. Ce problème a ensuite été étendu à la fouille de séquence, permettant ainsi des applications dans la génomique ou pour l'extraction de formes temporelles, par exemple dans des réseaux de télécommunications. Récemment, le problème a été étendu à des données encore plus complexes comme la fouille d'arbres fréquents, et plus généralement de sous-graphes fréquents pour des applications telles que la biochimie, la fouille dans le world wide web, les réseaux de citation, et l'analyse de flux de données dans les programmes. Nous proposons dans ce chapitre une introduction à ce domaine, et particulièrement à l'algorithme ECLAT que nous étendrons au cas particulier des STIS au ch.9. Le lecteur intéressé trouvera dans l'annexe ?? une présentation plus détaillées du domaine et des différents algorithmes utilisés.

6.1 Problème et définitions

Soit $\mathcal{S} = \{s_1, \dots, s_{|\mathcal{S}|}\}$ un ensemble de $|\mathcal{S}|$ symboles.

Soit \mathcal{O} un ensemble d'objets o_i de longueurs variables : $\mathcal{O} = \{o_1, \dots, o_{|\mathcal{O}|}\}$. Chaque objet o_i contient un ensemble de k symboles $s_i \in \mathcal{S}$:

$$\forall o_i \in \mathcal{O}, o_i = \{s_i, s_i \in [1, k], s_i \in \mathcal{S}\} \quad (6.1)$$

Un tel objet peut représenter un mot constitué de lettres, ou un ensemble de nucléotides d'une séquence d'ADN, ou encore des transactions d'achat. Dans ce dernier cas, les objets ne contiennent pas de répétition. On a donc $o_i \subseteq \mathcal{S}$, et c'est ce cas particulier que nous étudions par la suite. La figure 6.1 montre quelques exemples d'objets. On indexe chaque objet par son identifiant, i . On note I l'ensemble des identifiants d'objets. On appelle motif, et on note m , un ensemble de symboles. Les symboles d'un motif peuvent être ordonnés par

1. item set mining

i	o_i
1	$\{a, d, e\}$
2	$\{b, c, d\}$
3	$\{a, c, e\}$
4	$\{a, c, d, e\}$
5	$\{a, e\}$
6	$\{a, c, d\}$
7	$\{b, c\}$
8	$\{a, c, d, e\}$
9	$\{b, c, e\}$
10	$\{a, d, e\}$

FIGURE 6.1 – Exemple de base de données. La première colonne représente les identifiants des objets, et la deuxième les objets correspondants.

un ordre total (lexicographique par exemple), et on note alors la succession de symboles de $m : m[1].m[2]....m[k]$. Le nombre de symboles, k de m est appelé longueur de m : $k = l(m)$. On parle dans ce cas de k -motif. On appelle n -extension de m_1 le k -motif m_2 si et seulement si $m_2 = m_1.m_3$ où m_3 est un n -motif. m_1 est alors le $(k-m)$ -préfixe de m_2 .

On dit que l'objet o contient (ou couvre) le motif m si et seulement si $m \subseteq o$, pour deux motifs m_k et m_{k+1} , si $m_k \subset m_{k+1}$ alors on dit que m_{k+1} est un motif couvrant de m_k . La fréquence, ², σ d'un k -motif m dans la base de données est le nombre d'objets qui contiennent m . On dit qu'un k -motif est fréquent si il est contenu dans au moins σ_{\min} objets, où σ_{\min} est un paramètre de l'algorithme. L'ensemble des motifs fréquents est noté \mathcal{F} et l'ensemble des k -motifs fréquents est noté \mathcal{F}_k .

$$\mathcal{F}_k = \{m \in \mathcal{S} \mid l(m) = k, \text{ et } \sigma(m) \geq \sigma_{\min}\} \quad (6.2)$$

Un exemple d'une base de données est présenté figure 6.1.

La couverture K d'un motif m est l'ensemble des identifiants des objets couvrants de m : $K(f) = \{j \in \{1, \dots, |\mathcal{O}|\} \mid f \subseteq o_j\}$

6.2 Propriétés générales

Trouver les motifs fréquents en énumérant simplement tous les motifs présents dans \mathcal{O} , puis en calculant leur fréquence de façon à supprimer les non fréquents est généralement impossible en raison de la complexité de cette approche : $O(2^{|\mathcal{S}|})$.

Les algorithmes de recherche de motifs fréquents exploitent donc des propriétés de la fréquence des motifs de façon à ne calculer la fréquence que pour un nombre réduit de motifs. Dans un algorithme de recherche de motifs fréquents, les motifs dont la fréquence est calculée sont appelés *candidats*. Idéalement, le nombre de candidats est égal au nombre de motifs fréquents, cependant, un tel algorithme n'existe pas, et l'objectif est donc de réduire l'ensemble des candidats de façon à réduire le temps CPU tout en faisant face à des contraintes de stockage.

2. Dans la littérature, on parle aussi de support, mais pour plus de clarté dans le chapitre 9 où nous réutiliserons ces méthodes, et où on manipulera un autre type de support (temporel), nous avons préféré le terme fréquence. Ce terme est aussi utilisé dans la littérature, normalisé par rapport à $|\mathcal{O}|$.

Remarquons que la fréquence d'un motif ne peut que décroître quand il est étendu. Cette propriété, appelée *propriété d'antimonotonie* de l'ensemble des symboles, ou encore *propriété de fermeture vers le bas* (car l'ensemble des motifs fréquents est fermé par l'inclusion) est utilisée par tous les algorithmes de recherche d'ensembles fréquents de façon à réduire l'espace de recherche :

$$\forall m : \forall m' \supseteq m \Rightarrow \sigma(m') \leq \sigma(m) \quad (6.3)$$

En conséquence :

Propriété 6.2.1 : *Aucun motif couvrant d'un motif non-fréquent n'est fréquent.*

Cette propriété dite Apriori, sur laquelle s'appuient tous les algorithmes, permet de réduire beaucoup l'espace de recherche dans une approche par croissance de motif.

Réciproquement, les sous-motifs d'un motif fréquent sont aussi fréquents.

Il existe trois types de motifs fréquents :

- Motif (fréquent) libre, $m : \sigma(m) > \sigma_{\min}$. Cet ensemble est noté L .
- Motif (fréquent) fermé, m : Il n'existe pas de motif couvrant de m de même fréquence. Cet ensemble est noté C .
- Motif (fréquent) maximal, m : Il n'existe pas de motif couvrant de m fréquent. (les motifs maximaux sont aussi fermés). Cet ensemble, noté M est parfois aussi appelé bordure positive.

Propriété 6.2.2 : *Tout motif fréquent a un motif couvrant maximal.*

Propriété 6.2.3 : *Tout motif fréquent est un sous-ensemble d'un motif maximal*

En conséquence, l'ensemble des motifs maximaux est une représentation condensée de l'ensemble des motifs fréquents. Et elle est bornée par $\binom{|\mathcal{S}|}{|\mathcal{S}|/2}$ d'après le théorème de Sperner's. Nombreuses sont les applications où l'ensemble des motifs fréquents est très grand. En particulier pour des données contenant des longs motifs fréquents car dans ce cas chaque motif maximal a un grand nombre de sous-motifs. Le nombre de motifs fréquents croît d'ailleurs exponentiellement avec la longueur. En effet, un motif fréquent de taille k inclut $2^k - 1$ sous-motifs fréquents. Au ch.9, c'est cet ensemble que nous chercherons à extraire.

Cependant, l'ensemble des motifs fréquents maximaux ne permet pas de renseigner sur la valeur de la fréquence des motifs fréquents qu'il couvre. L'ensemble des motifs fermés, en revanche, permet de disposer de cette information. C'est pourquoi cette forme condensée est beaucoup utilisée pour la génération de règles pour lesquelles la connaissance des fréquences est nécessaire.

En effet, d'après la définition des motifs fréquents fermés, un motif fréquent est soit fermé, soit il dispose d'un motif couvrant de même fréquence. Une autre caractérisation des motifs fermés est qu'ils sont égaux à l'intersection de tous les objets de \mathcal{O} les couvrant.

La propriété Apriori relie les motifs avec leurs sous-motifs et leurs motifs couvrants. En conséquence, le treillis des sous-ensemble de motifs paraît particulièrement adapté pour organiser la recherche. L'ensemble vide est l'élément du haut et l'ensemble de tous les symboles est l'élément du bas d'un tel treillis. Ce treillis est représenté pour quatre symboles figure 6.2. Les approches par croissance de motif exploitent l'arbre de préfixe contenu dans le treillis. (voir figure 6.2b). Il existe deux approches pour parcourir cet arbre : un parcours

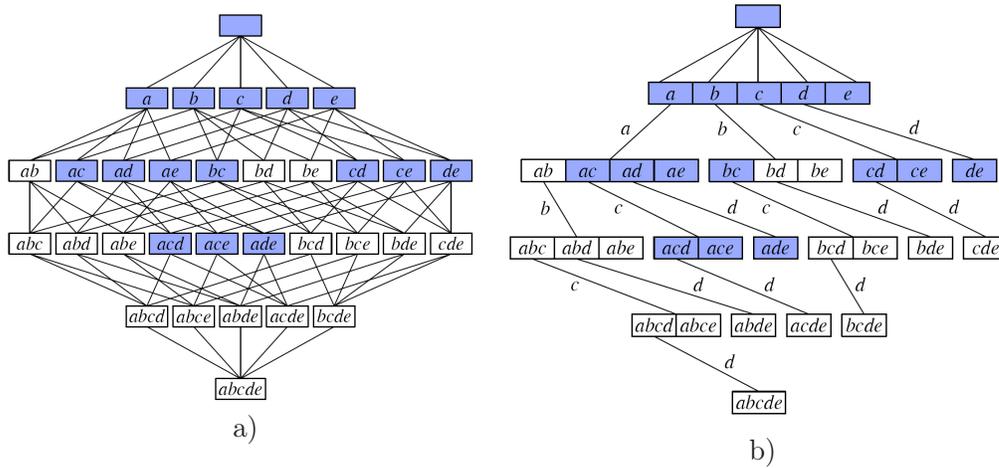


FIGURE 6.2 – a) : Diagramme de Hasse : Treillis pour cinq symboles. En bleu, les motifs fréquents correspondant à la base de données présentée figure 6.1. b) : Arbre de préfixe correspondant à la base de données présentée figure 6.1

en largeur, et un parcours en profondeur. Nous allons tout d'abord présenter succinctement l'algorithme Apriori qui suppose un parcours en largeur, puis nous présenterons plus en détail l'algorithme ECLAT qui consiste en un parcours en profondeur car c'est cette approche que nous avons adopté par la suite.

6.3 ECLAT

L'algorithme ECLAT est le premier algorithme proposé pour la recherche de motifs fréquents par un parcours en profondeur de l'arbre de préfixe [Zaki, 2000c, Zaki et al., 1997]. L'ensemble complet des sous-motifs d'un candidat n'est pas connu lors d'un parcours en profondeur, et la propriété d'antimonotonie ne peut donc pas être complètement exploitée. En effet, seule l'étape de jointure des motifs de la génération de candidats est effectuée. Il en résulte un nombre plus important de candidats. Cependant, une représentation verticale de la base de données, et un calcul de la fréquence par intersection des lignes de cette représentation verticale (comme dans [Savasere et al., 1995]) permet un calcul rapide de leur fréquence, et par conséquent, le coût CPU n'est pas celui qu'on pourrait attendre, au contraire, il est souvent plus faible que par des algorithmes de type Apriori.

ECLAT génère récursivement les ensembles conditionnels pour chaque symbole $s \in \mathcal{S}$. L'algorithme est présenté figure 6.3.

Chaque 1-motif, m , constitué d'un symbole fréquent s est agrandi par extension pour tous les 1-motif fréquent plus grands : $r > m[0]$. La couverture de cette extension est alors calculée par intersection de leur couvertures. Si la fréquence associée, c'est-à-dire le cardinal de leur couverture est supérieure au seuil σ_{\min} , alors, r est ajoutée à l'ensemble conditionnel d'objets,³ de s : \mathcal{O}_s . Puis, on réitère ces étapes pour chacun des éléments de cet ensemble conditionnel d'objets, et ainsi de suite, jusqu'à ce que plus aucune extension ne soit fréquente.

Un réordonnement des symboles par fréquence croissante permet de réduire le nombre de candidats. Un autre moyen de réduire le temps de calcul des fréquences, est d'utiliser

3. Dans la littérature, on parle de base de données conditionnelle

```

Entrées :  $\mathcal{O}, \sigma_{\min}, \mathcal{S};$ 
Sorties :  $\mathcal{F};$ 
1 début
2    $\mathcal{F} \leftarrow \emptyset;$ 
3   pour chaque  $s \in \mathcal{S}$  faire
4     | si  $\sigma(s) > \sigma_{\min}$  alors  $\mathcal{F} \leftarrow \mathcal{F} \cup s$ 
5   fin
6   pour chaque  $m \in \mathcal{F}$  faire
7     | pour chaque  $r \in \mathcal{S} | \sigma(r) > \sigma_{\min} \ \& \ r > m[l(m) - 1]$  faire
8       | |  $C \leftarrow K(\{s\}) \cap K(\{r\});$ 
9       | | si  $|C| \geq \sigma_{\min}$  alors  $\mathcal{F} \leftarrow \mathcal{F} \cup \{m.r\}$ 
10    | fin
11  | fin
12 fin

```

FIGURE 6.3 – Algorithme ECLAT. On étend chaque 1-motif fréquent récursivement en lui ajoutant un symbole fréquent plus grand que le dernier symbole du motif : $m[l(m) - 1]$, où $l(m)$ est la longueur du motif.

	1	2	3	4	5	6	7	8	9	10
A	1	0	1	1	1	1	0	1	0	1
B	0	1	0	0	0	0	1	0	1	0
C	0	1	1	1	0	1	1	1	1	0
D	0	1	0	1	0	1	0	1	0	1
E	1	0	1	1	1	0	0	1	1	1

FIGURE 6.4 – Représentation verticale de la base de données de la figure 6.1

un autre formatage des données appelé *format vertical*. [Savasere et al., 1995, Dunkel and Soparkar, 1999, Shenoy et al., 2000, Zaki, 2000c]. En effet, la fréquence d'un ensemble de symboles est le cardinal de son ensemble d'identifiants d'échantillons. Or ces derniers se calculent par de simples intersections. De plus, ces intersections permettent une réduction de l'information non pertinente : seules les identifiants des échantillons nécessaires au calcul de la fréquence restent après une intersection.

Si l'on reprend l'exemple de la figure 6.1, la représentation verticale correspondante est présentée figure 6.4.

En supposant que $\sigma_{\min} = 2$, l'algorithme ECLAT commence par générer la base de données conditionnelle de A. C'est-à-dire la représentation verticale de AB, AC, AD, et AE, qui est obtenue par simple & logique du vecteur de bits correspondant à chacun des symboles. Seuls AC, AD, et AE sont fréquents. On les étend alors à un autre symbole plus grand que le dernier symbole du motif : ACD et ADE. Et finalement, ACDE. La figure ?? illustre la génération de la base de données conditionnelle de A.

[Zaki and Gouda, 2003] a proposé un stockage plus efficace des ensembles conditionnels d'objets par *diffsets*. Il s'agit des différence de couvertures entre le k-motif s , et son k-1-préfixe. [Hipp et al., 2000] propose de combiner Apriori et ECLAT dans un algorithme appelé Hybride. L'algorithme commence la génération des motifs fréquents par une approche Apriori, puis, pour une certaine profondeur, fixée par l'utilisateur, l'algorithme

	1	2	3	4	5	6	7	8	9	10
AC	0	0	1	1	0	1	0	1	0	0
AD	0	0	0	1	0	1	0	1	0	1
AE	0	0	1	1	1	0	0	1	0	1
ACD	0	0	0	1	0	1	0	1	0	0
ADE	0	0	0	1	0	0	0	1	0	1
ACDE	0	0	0	1	0	0	0	1	0	0

FIGURE 6.5 – Génération de la base de données conditionnelle de A pour l'algorithme ECLAT pour les données de la figure 6.4.

passé à une génération en profondeur façon ECLAT.

Quatrième partie

Mise en oeuvre et implémentation

Chapitre 7

Segmentation jointe d'une STIS

Comme nous l'avons vu au chapitre 4, les STIS peuvent être modélisées dans un graphe d'adjacence temporelle d'objets spatiaux (SOTAG). Les objets spatio-temporels (STO) qui sont des séquences temporelle des ces objets spatiaux (SO) s'obtiennent à partir des chemins dans une version orientée de ce graphe. Les SO sont caractérisés par une homogénéité spatiale radiométrique. Par ailleurs, les régions des STO sont modifiées par des événements spatiaux de type fusion, division, et la combinaison des deux. Il existe donc une homogénéité temporelle de forme entre groupes d'SO de différents instants. Cette homogénéité est néanmoins perturbée par de faibles déformations liées à des erreurs de recalages géométriques entre images ou à de faibles déformations des STO de la scène dont on a sous-estimé la profondeur.

On peut envisager de traiter ces deux types de redondances (spatiale et temporelle) successivement, par exemple en segmentant chacune des images de la STIS, puis en mettant en correspondance les objets le mieux possible de façon à mettre en évidence les STO et les changements topologiques qui les lient. Cependant, en raison des changements radiométriques liés aux conditions d'acquisition ou aux objets eux-même, les objets spatiaux extraits indépendamment pour chaque image par un algorithme de segmentation classique présentent généralement peu de redondance de forme. En effet, pour pouvoir parler de forme d'un objet, il faut disposer de l'objet, mais pour assurer que ces objets sont effectivement des instances d'un SO pertinent, il faut qu'ils présentent une forte redondance de forme. Nous voyons donc que ces deux redondances caractérisant les STO sont très liées et nous devons donc les traiter conjointement. Nous proposons donc une méthode de segmentation jointe de STIS.

7.1 Introduction

Les régions spatiales que l'on recherche subissent par ordre de fréquence les changements suivants :

- Une conservation topologique : Correspondance 1-1
- Une fusion ou une division : Correspondance 1-n
- Une combinaison de fusion(s) et de division(s) : Correspondance n-m

Nous disposons donc d'un modèle de la séquence que l'on peut voir comme un *langage* permettant de décrire la séquence. Restent à mettre en évidence les régions spatiales dont les évolutions se décrivent simplement à l'aide de ce langage.

Remarquons que les changements les plus rares sont les plus compliqués : une combinaison de fusion(s) et de division(s) correspondant à un appariement n - m est nettement plus compliqué qu'une fusion ou qu'une division qui correspondent à des appariements 1 - n et qui sont eux-mêmes plus compliqués qu'une conservation correspondant à un appariement 1 - 1 . Par ailleurs, l'objectif de la segmentation est de simplifier les données, ce qui est une problématique proche de celle du codage, et comme nous l'avons présenté dans 5, formuler le problème de la segmentation d'image sous la forme d'un problème de codage permet de mettre en évidence une énergie minimale pour la partition optimale correspondant à l'image. Par ailleurs, cette formulation permet une mise à échelle des différents types d'information manipulés en raison de l'unité de mesure universelle considérée : le bit. Ce point est particulièrement intéressant dans notre cas où nous devons traiter conjointement deux redondances de nature différentes : radiométrique et de forme. Nous avons donc choisi un cadre de codage pour la segmentation jointe de la STIS, et particulièrement le cadre du MDL.

Nous avons introduit dans la section A.7 le SOTAG. Nous introduisons maintenant le graphe d'adjacence spatio-temporelle de régions (SOSTAG, Spatial Object Spatio-Temporal Adjacency Graph) car c'est une structure permettant de mettre en évidence les relations topologiques spatiales (utiles à un algorithme de segmentation) et temporelles. De plus, nous verrons §7.5 que ce graphe attribué est une structure de stockage des paramètres nécessaires à l'optimisation permettant une mise à jour efficace.

Définition 7.1.0.1 SOSTAG: *Un SOTAG est un graphe dont les nœuds représentent les régions spatiales, et dont les arcs sont de deux types :*

- *spatiaux : deux nœuds sont liés par un arc spatial si les régions associées ont un contour commun.*
- *temporels : deux nœuds sont liés par un arc temporel si les régions associées ont une intersection non vide.*

7.2 Modèle géométrique spatio-temporel

Dans cette section, nous présentons le schéma de codage de la séquence. Nous commençons par présenter le schéma de codage idéal, et après en avoir discuté la complexité, nous présentons les approximations que nous avons considéré.

7.2.1 Codage différentiel de la géométrie

Nous avons vu en A.2.3.3 qu'il est classique en codage vidéo, de décrire une série de partitions à partir d'une partition initiale, codée en intra, et la série des modifications des régions entre chaque nouvelle trame à partir de la compensée [Salembier et al., 1997, Reusens, 1995]. Cette approche différentielle est plus intéressante d'un point de vue de codage car le nombre de bits requis pour coder les modifications, et les erreurs résultantes est inférieur au nombre de bits requis pour coder l'ensemble des partitions en intra. Ce gain est dû d'une part à l'existence d'une redondance temporelle dans les données codées, et d'autre part à l'exploitation de cette redondance dans le schéma de codage. Dans le cas d'une séquence vidéo, cette redondance temporelle est radiométrique et géométrique (voir). Dans le cas des STIS, nous ne considérons pas de redondance radiométrique en raison des évolutions radiométriques des objets de la scène ainsi que des changements radiométriques liés aux changements des conditions d'acquisition. En revanche, la redondance géométrique

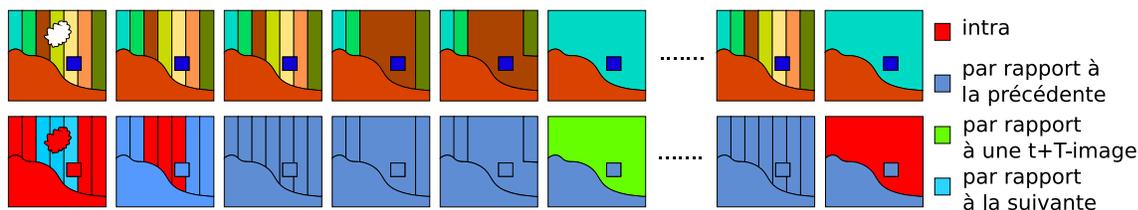


FIGURE 7.1 – Schéma de codage idéal. Dans un schéma de codage idéal, pour chaque région, il faut choisir une image de référence et le type de changement à considérer.

est importante, et il est donc intéressant de l'exploiter. Nous proposons donc un codage différentiel de la géométrie des régions.

7.2.2 Image de référence

Dans le cas de la vidéo, les changements sont décrits d'une trame à l'autre de façon transitive selon l'ordre chronologique. Or nous avons vu que dans notre cas, un nuage peut apparaître dans une image. Dans ce cas, dans l'image suivant l'image du nuage, il sera moins coûteux de décrire les régions désoccultées par rapport à l'avant-dernière image. Par ailleurs, l'un des avantages des STIS par rapport aux paires d'images utilisées pour détecter des changements est la durée pendant laquelle la terre est observée. Cette longue durée, qui permet d'observer l'histoire des objets est particulièrement intéressante dans le cas de phénomènes périodiques. Dans ce cas, une mémoire de la taille de la période P du phénomène observé est requise car il est plus intéressant de coder les régions de l'instant t par rapport à une image de l'instant $t - k.P$. Dans ce cas, les régions sont donc codées de façon plus efficace par rapport à des régions plus anciennes que la précédente voire même futures. La figure 7.1 illustre ce codage pour un exemple jouet.

7.2.3 Différents types de changements

En raison des erreurs éventuelles de recalage géométrique, des superpositions, ou de segmentation (particulièrement si l'on considère une méthode itérative), les correspondances entre régions spatiales ne correspondent pas à une simple intersection des régions. Il nous faudra donc déterminer, pour chaque région, et par rapport à chacune des autres images, le type de changement à considérer, puis, déterminer l'image par rapport à laquelle cette région doit être décrite, c'est-à-dire l'image permettant la description du changement la plus simple.

7.2.4 Description idéale et complexité

La description idéale de la géométrie de la STIS consiste en une description conditionnelle. Certaines régions-clés sont codées en intra de façon à permettre la description d'autres régions par rapport à celles-là de façon conditionnelle. Le choix des régions clés et des images de référence par rapport auxquelles décrire les régions en inter est cependant un choix difficile. En effet, l'image de référence associée à une région à coder en inter doit permettre la description globale la plus simple possible. Ce choix fait intervenir dans un premier temps la complexité de description conditionnelle par rapport à chacune des images, mais est aussi lié au choix des régions clés et au choix des autres dépendances qu'il entraîne car pour éviter les redondances, une région doit être décrite par rapport à une

seule image. Ainsi, dans l'exemple de la figure 7.1, il faut s'assurer qu'au bout de la chaîne de dépendance de la région r_{10} il y a bien une région codée en intra, c'est-à-dire s'assurer qu'il n'y a pas de cycles, et que le choix de codage en intra est judicieux pour cette région.

Par ailleurs, la description conditionnelle d'une région par rapport à une certaine image de référence nécessite d'avoir au préalable déterminé le type de changement à considérer. La détermination du type de changement nécessite de tester chacune des quatre hypothèses de changement. L'hypothèse de conservation est assez simple à tester : la région de l'image dans laquelle on cherche les correspondances permettant la meilleure correspondance est sélectionnée. La complexité de cette étape est donc de l'ordre du nombre moyen de voisins temporels d'une région. Et la longueur de description géométrique sous cette hypothèse revient à coder le numéro de la région en question, ceux des régions à partir desquelles elle est codée, et les modifications à effectuer pour passer d'une région à l'autre correspondant à l'erreur géométrique d'approximation. Le cas d'une fusion ou d'une division est plus complexe : en fonction de la taille de la région fusionnée (respectivement, qui se divise), on peut avoir un nombre plus ou moins grand de combinaisons de régions avec lesquelles tester la mise en correspondance. Ce problème revient en effet à chercher, dans un graphe d'adjacence spatio-temporelle de régions, la composante spatialement connexe dans l'image de référence considérée présentant la meilleure correspondance parmi toutes les composantes spatialement connexes dont tous les nœuds sont temporellement liés à la région fusionnée (respectivement à diviser). Une fois le groupe optimal obtenu, coder la région revient à coder les numéros de régions correspondantes, et les modifications à effectuer sur la réunion pour obtenir l'autre région. Ces modifications permettent la robustesse aux petites déformations. Enfin, trouver le nombre de mises en correspondance n-m est encore plus complexe car cela revient à trouver des paires de composantes spatialement connexes dont chaque nœud est temporellement lié à un nœud de l'autre composante spatialement connexe. Enfin, cette optimisation doit être effectuée pour chacune des régions.

Nous voyons donc que la description de la séquence de façon différentielle selon ce schéma de codage est extrêmement complexe. De plus, dans le cas d'un algorithme itératif, les différentes étapes de décision du type de changement, de choix de l'image référence, et de recherche des région-clés doivent être répétées à chaque itération. L'optimisation jointe des appariements optimaux et des segmentations spatiales n'est donc pas envisageable pour ce modèle sans effectuer d'approximation. Nous proposons donc deux approximations.

7.2.5 Décision concernant le choix du type de changements

Nous avons vu que le choix parmi les quatre types de changements est très coûteux, et particulièrement les combinaisons de fusion(s) et de division(s). Notons qu'une combinaison de fusion(s) et de division(s) peut aussi être décrite en termes de fusions, et de divisions, et les fusions et les divisions peuvent aussi être décrites en termes de conservation topologique. Néanmoins, dans ces cas, les descriptions n'étant pas adaptées aux données, les erreurs géométriques résultantes seront plus importantes. Par ailleurs, les régions spatiales des STIS que nous considérons sont généralement conservées.

Afin de simplifier le problème, nous proposons donc de considérer uniquement des correspondances 1-1, quitte à coder des erreurs géométriques importantes. En conséquence, étant donné que les STIS contiennent des changements topologiques, une région pourra être référence pour plusieurs régions, et certaines régions ne seront référence d'aucune région. La conséquence de cette approximation est que le coût de description d'un autre changement que la conservation topologique est surestimé. En effet, cette description est très redondante

dans le cas d'une division par exemple, parce qu'on décrit la transformation de chacune des régions divisées vers la région qui se divise. La description est encore plus redondante dans le cas d'une combinaison de ces changements. L'ordre de complexité initial est néanmoins respecté : la longueur de description d'une conservation topologique est inférieure à la longueur de description d'une fusion ou d'une division elle-même inférieure à la longueur de description d'une combinaison.

La redondance de cette description n'est pas gênante en soi dans la mesure où l'on ne souhaite pas coder effectivement la séquence. Elle nous écarte cependant du cadre formel de codage et la longueur de description doit être vue plutôt comme une fonctionnelle à optimiser. Cette redondance a aussi pour conséquence d'affecter la commensurabilité des différents termes de la longueur de description. La partie radiométrique perd ainsi de l'importance relativement à la partie géométrique dans des cas de changements topologiques. Ceci aura pour conséquence une adaptation de la sensibilité radiométrique par rapport à la redondance temporelle.

7.2.6 Décision concernant le choix de l'image de référence

Le modèle présenté contient des arcs entre images non consécutives afin d'être robuste aux apparitions de nuages et de pouvoir exploiter la redondance temporelle de phénomènes périodiques. Cette approche nécessite le choix de l'image référence pour chaque région, et nécessite ensuite de rompre les cycles présents dans le graphe de dépendance en codant l'une des régions en intra afin d'assurer le décodage. Ainsi, plus le graphe contiendra de cycles, plus il faudra coder de régions en intra. Donc plus les chemins maximaux sans cycles sont longs, moins il faudra coder de régions en intra. Séparer cette décision du choix de l'image de référence est sous-optimal. En effet, choisir l'image de référence d'une région comme étant celle qui présente la modification la plus simple à décrire peut conduire à un petit chemin maximal sans cycle alors qu'en prenant une autre image, on aurait pu obtenir un chemin maximal sans cycle plus grand, et donc coder, au total, moins de régions en intra. Une optimisation globale est impossible en raison du nombre de configurations à tester, néanmoins, différentes approximations peuvent être effectuées de façon à trouver un optimum local. Rappelons cependant que la mise en correspondance n'est pas l'objectif final de la méthode de segmentation jointe. Les régions de chaque image sont initialement inconnues. Cette mise en correspondance interviendra donc à chaque itération d'une autre optimisation et doit donc être la plus simple possible.

Afin de régler le problème de cycles, nous proposons d'imposer un codage par rapport au passé. Cette décision correspond à une vision classique correspondant à un temps orienté.

Le choix de la région de référence passée reste néanmoins entier, et nous proposons d'éviter cette prise de décision coûteuse en considérant une moyenne des longueurs de description correspondant à chaque configuration. La longueur de description géométrique d'une région r s'écrit alors :

$$\mathcal{L}_{\text{geom}}^{\text{ST}}(r) = \frac{1}{|\mathcal{C}_{\text{yl}}(r)| + 1} \left(\mathcal{L}_{\text{geom}}^{\text{TL}}(r) + \sum_{r' \in \mathcal{C}_{\text{yl}}(r)} \mathcal{L}_{\text{geom}}^{\text{TD}}(r|r') \right) \quad (7.1)$$

où $\mathcal{C}_{\text{yl}}(r)$ est l'ensemble des régions correspondant à r en chaque autre instant. Cette longueur de description favorise donc effectivement les régions stables car dans ce cas, le terme $\mathcal{L}_{\text{geom}}(r|r')$ est nul (ou quasiment dans le cas d'erreur de recalage géométrique de la région). En revanche, les images avec lesquelles la correspondance est mauvaise seront aussi prises en compte.

Outre le problème de complexité de l'optimisation, dans le cas d'images mal recalées ou dans le cas d'erreurs dans les partitions considérées (ce qui arrivera au cours d'un algorithme itératif) le graphe intersection à considérer pour calculer ces dépendances peut vite comprendre un nombre prohibitif d'arcs. Ainsi, il n'est pas envisageable de calculer le graphe d'intersection complet pour la séquence ADAM entière ou même d'une sous-séquence obtenue par découpage spatial. On peut alors simplifier encore le problème en réduisant le nombre d'images candidates à la dépendance. Autrement dit, on peut réduire l'influence d'une image à une fenêtre temporelle. L'intérêt des arcs entre images non consécutives était de prendre en compte les phénomènes périodiques et les nuages. On peut donc réduire les images candidates à la dépendance à :

- L'image précédente
- L'image avant la précédente; afin de rester robustes aux apparitions liées à une superposition
- Les images à des multiples de un an d'intervalle; dans ce cas, la notion de passé ou de futur perd son sens puisqu'on considère un temps circulaire.

Par la suite, on notera \mathcal{Cyl} l'ensemble des régions de référence, et W son cardinal.

Enfin, l'approximation la plus forte consiste à prendre $W = 1$, et ne considérer de dépendance qu'à l'image précédente. Dans ce cas, le problème du choix de l'image de dépendance n'est plus. En revanche, les apparitions liées à une superposition comme les nuages auront pour effet de considérer les régions occultées comme instables à l'instant suivant, et donc soit de sursegmenter les deux soit de sous-segmenter. En effet, on remarque que le cas de la sur-partition triviale où chaque pixel est une région correspond à une mise en correspondance parfaite. Dans le cas où la réalité présente de mauvaises mises en correspondances, on pourra rester dans ce minimum local de la longueur de description spatio-temporelle.

7.3 Longueur de description du modèle géométrique spatio-temporel

De façon à calculer la longueur de description spatio-temporelle d'une région r_t (équation 7.1), il faut d'abord trouver $\mathcal{Cyl}(r_t)$, c'est-à-dire l'ensemble des régions (une par image de référence) correspondant le mieux à r_t . Il nous faudra ensuite calculer la longueur de description géométrique temporellement localisée $\mathcal{L}_{\text{geom}}^{\text{TL}}(r_t)$. Pour cela, nous renvoyons le lecteur à 5.2.1.1, et rappelons juste ici son expression :

$$\mathcal{L}_{\text{geom}}^{\text{TL}}(r_t) = \frac{1}{2} (\log(N) + \log(4) + (|\mathcal{C}(r_t)| - 1) \log(3)) \quad (7.2)$$

Finalement, il nous faudra calculer la description conditionnelle $\mathcal{L}_{\text{geom}}^{\text{TD}}(r_t)$. Nous détaillons dans cette section ces différentes étapes.

7.3.1 Calcul du cylindre d'une région

Le *cylindre* d'une région r_t est l'ensemble des régions des images de référence présentant la meilleure mise en correspondance avec r_t . Il existe de nombreuses mesures de mise en correspondance de forme. Remarquons que les images constituant la STIS ne sont pas tournées, et sont à même échelle, une variation d'échelle ou d'angle entre deux régions traduit donc un changement significatif. C'est pourquoi nous devons utiliser une mesure de mise en correspondance qui ne soit invariante ni par rotation, ni par changement d'échelle.

Il y a deux manières de comparer des régions : les comparaisons basées sur la région, et celles basées sur le contour. Nous allons présenter quatre mesures de distances.

Mesure de distance de frontière Cette distance, notée \mathcal{D}_B est la somme de deux mesures non symétriques $\mathcal{D}_B^d(\mathcal{C}_1 \rightarrow \mathcal{C}_2)$ et $\mathcal{D}_B^d(\mathcal{C}_2 \rightarrow \mathcal{C}_1)$ correspondant respectivement à la somme des distances de chaque élément d'un contour à l'autre contour :

$$\mathcal{D}_B = \mathcal{D}_B^d(\mathcal{C}_1 \rightarrow \mathcal{C}_2) + \mathcal{D}_B^d(\mathcal{C}_2 \rightarrow \mathcal{C}_1) \quad (7.3)$$

avec :

$$\mathcal{D}_B^d(\mathcal{C}_1 \rightarrow \mathcal{C}_2) = \sum_{c_1 \in \mathcal{C}_1} d(c_1, \mathcal{C}_2) \quad (7.4)$$

où :

$$d(c_1, \mathcal{C}_2) = \min_{c_2 \in \mathcal{C}_2} d(c_1, c_2) \quad (7.5)$$

où $d(c_1, c_2)$ représente la distance euclidienne entre deux éléments de contour. Cette mesure considère équidistantes d'une région r_1 deux régions r_2 et r_3 , r_2 présentant, par rapport à r_1 de faibles erreurs locales entre les contours tout le long du contour (par exemple suite à une légère translation), et r_3 présentant des erreurs importantes, mais sur une petite partie du contour.

Distance de Hausdorff La distance de Hausdorff, \mathcal{D}_H consiste à prendre le maximum, plutôt que la somme par rapport à la distance de frontières.

$$\mathcal{D}_H = \max(\mathcal{D}_H^d(\mathcal{C}_1 \rightarrow \mathcal{C}_2), \mathcal{D}_H^d(\mathcal{C}_2 \rightarrow \mathcal{C}_1)) \quad (7.6)$$

où :

$$\mathcal{D}_H^d(\mathcal{C}_1 \rightarrow \mathcal{C}_2) = \max_{c_1 \in \mathcal{C}_1} (d(c_1, \mathcal{C}_2)) \quad (7.7)$$

Cette distance quantifie la pire différence entre deux contours et est donc bien adaptée à la mise en correspondance de formes. En reprenant l'exemple précédant, cette mesure place r_3 plus loin de r_1 que r_2 .

Distance de Hamming Définie par Richard Hamming, cette distance joue un rôle très important en théorie algébrique. Elle permet de quantifier la différence entre deux séquences de symboles. Si on pose $a = a_0 \dots a_{n-1}$ et $b = b_0 \dots b_{n-1}$, où les a_i et b_i sont des symboles appartenant à un ensemble A, la distance de Hamming se définit formellement par $d(a, b) = |\{i : a_i \neq b_i\}|$. Elle peut être étendue à des régions [Kanungo et al., 1994]. Elle est dans ce cas une distance basée région qui mesure l'aire des parties de régions n'étant pas commune aux deux régions dont on cherche la distance, mais appartenant simplement à l'une.

$$\mathcal{D}_{Ham}(r_1, r_2) = |r_1| + |r_2| - 2 \cdot |r_1 \cap r_2| \quad (7.8)$$

Distance logarithmique de régions On peut aussi considérer une version logarithmique de la distance de Hamming, \mathcal{D}_R en considérant le logarithme du rapport du produit des surfaces des régions et du cardinal de leur intersection.

$$\mathcal{D}_R(r_1, r_2) = \log(|r_1|) + \log(|r_2|) - 2 \cdot \log(|r_1 \cap r_2|) \quad (7.9)$$

Discussion Pour la mesure de distance de frontière, comme pour la distance de Hausdorff, le calcul de la distance nécessite pour chaque élément de contour, le calcul de la distance à tous les éléments de contours de toutes les régions ayant une intersection non vide. Soit \overline{N}_n le nombre moyen de régions ayant une intersection non vide, \overline{k} le nombre moyen de régions de l'image, et $N_x * N_y$ la taille de l'image. On peut approximer le nombre moyen de contours par région par $2\sqrt{N_x * N_y / \overline{k}}$. En conséquence, la complexité du calcul de cette distance pour une paire d'images est $4\overline{N}_n N_x * N_y$, et la mise à jour de cette distance pour une région est $4\overline{N}_n N_x * N_y / \overline{k}$. Pour la distance de région, en revanche, la complexité d'initialisation des intersections est en $O(N_x * N_y)$ car un simple parcours de l'image et une incrémentation des compteurs dans un tableau de taille \overline{k}^2 suffit et la mise à jour d'une distance est en temps constant. Nous verrons par la suite qu'un algorithme itératif de fusions spatiales des régions est bien adapté pour minimiser la longueur de description spatio-temporelle, et on préférera donc une mesure basée région. En particulier, nous choisirons la distance logarithmique de régions.

7.3.2 Longueur de description géométrique conditionnelle

La description considérée pour les régions est une description du contour. Décrire une région à partir d'une autre revient donc à décrire un contour à partir d'un autre contour. La façon la plus simple de le faire est de décrire les portions de contour à enlever, et les portions de contour avec lesquelles les remplacer. Notons que connaissant l'ancien contour, et les portions du nouveau contour à insérer, on peut retrouver les portions du contour initial à remplacer. Il s'agit de la portion de l'ancien contour situé entre les deux extrémités du contour à insérer. Seules les nouvelles portions de contour doivent donc être décrites.

Si les contours sont complètement disjoints, on devra décrire totalement le nouveau contour. La description conditionnelle permettra une économie de longueur de description par rapport à la description indépendante de la nouvelle région seulement dans le cas où il existe des portions de contour communes. Cette description est donc très sensible aux erreurs de recalage. On souhaiterait décrire les portions du nouveau contour à remplacer seulement si leur description est vraiment nécessaire, autrement dit si l'erreur commise si on garde les anciennes portions à la place n'est pas trop grande. Pour ce faire nous proposons de considérer une distance minimale entre deux éléments de contour parallèles l'un de l'ancien contour, et l'autre du nouveau contour. Seuls les segments du contour à insérer dont la distance à l'ancien contour est supérieure à cette distance minimale sont codés.

Enfin, remarquons que la longueur de description conditionnelle n'est pas symétrique. Cette asymétrie a pour effet de favoriser les rétrécissements de régions par rapport aux expansions. Cet effet est compensé par le fait que les régions voisines sont en contrepartie étendues ou rétrécies. Cependant, étant donné qu'on ne calcule cette longueur que sur des cylindres, c'est-à-dire sans prendre en compte les voisins spatiaux, cette asymétrie est gênante. Nous proposons donc de considérer une version symétrisée de la longueur de description conditionnelle en codant la moitié de la somme des longueurs de description du contour à remplacer et du contour à insérer.

Considérons un ensemble de segments constitué de la réunion des segments des contours extérieurs des deux régions dont on cherche la longueur de description conditionnelle. Les contours à décrire sont les cycles dont l'intérieur n'appartient pas à l'intersection de régions. La figure 7.3 illustre les contours à coder que l'on appellera *contours d'erreur*.

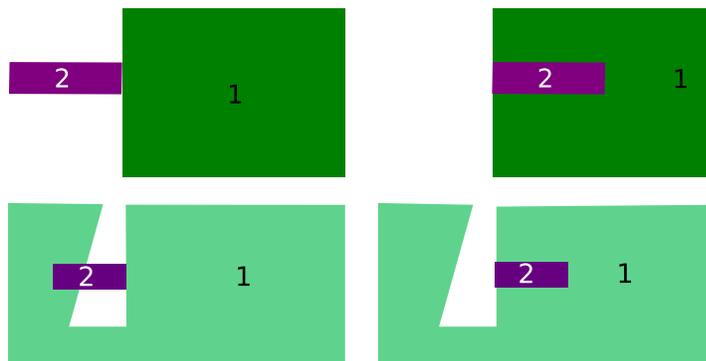


FIGURE 7.2 – Illustration de l'intérêt de considérer une orientation pour le contour.

Ces contours correspondent à plusieurs composantes connexes. On note \mathcal{C} l'ensemble de ces composantes connexes. Dans le cas de la figure 7.3), $\mathcal{C} = \{\mathcal{C}_1, \dots, \mathcal{C}_5\}$. La longueur de description associée est la moitié de la somme des longueurs des codes de chaîne de chacun de ces contours, le premier élément appartenant nécessairement au contour à partir duquel on souhaite le coder. Notons que ce facteur $1/2$ est à multiplier au facteur $1/2$ lié à ce qu'un contour est partagé par deux régions. La longueur de code de l'erreur de contour est donc :

$$\mathcal{L}_{\text{geom}}^{\text{TD}}(r_i|r_j) = \frac{1}{4} \left(|\mathcal{C}\mathcal{C}_{ij}| (\log(|\mathcal{C}_j|) + \log(4)) + \sum_{k=1}^{|\mathcal{C}\mathcal{C}_{ij}|} (|\mathcal{C}_k| - 1) \log(3) \right) \quad (7.10)$$

Ces composantes connexes sont donc constituées d'une portion du contour intérieur d'une des régions, et d'une portion du contour extérieur de l'autre région. La notion d'intérieur et extérieur d'un contour est prise en compte en considérant un sens conventionnel de parcours du contour dans le sens des aiguilles d'une montre du point de vue d'un point intérieur. Ce sens de parcours est nécessaire pour assurer qu'on ne compte pas comme contour commun une portion de contour dont l'une des régions est à gauche et l'autre à droite. Ainsi, dans l'exemple de la première ligne de la figure 7.2, aucun contour commun ne sera considéré entre les deux régions et donc les longueurs de descriptions conditionnelles reviennent simplement en la description complète du nouveau contour. A droite, en revanche, il existe un contour commun et la longueur de description conditionnelle sera donc la longueur de description du nouveau contour privé du contour commun. On ne calculera néanmoins jamais la longueur de description conditionnelle des régions de la figure de gauche car elles ont une intersection vide, et il est donc impossible que la région r_2 appartienne au cylindre de la région r_1 . Cependant, l'exemple de la ligne du dessous montre un cas pour lequel on pourra être amené à calculer la longueur de description conditionnelle, et alors, seul le cas de droite verra sa longueur de description conditionnelle réduite.

Les contours d'erreur forment donc des cycles dans l'ensemble des éléments de contour des régions, le sens de l'un d'eux ayant été inversé. Ainsi, dans le cas de la figure 7.3, les cycles présents dans l'ensemble des éléments de contour de chacun des contours dont le sens a été inversé pour l'un d'eux permet effectivement d'extraire les contours d'erreur. L'existence de points multiples complique néanmoins la recherche de cycles. Ainsi, supposons que l'on commence la recherche par le point vert. Arrivé au point bleu, deux possibilités sont envisageables. On choisira dans ce cas l'élément de contour permettant de fermer le plus

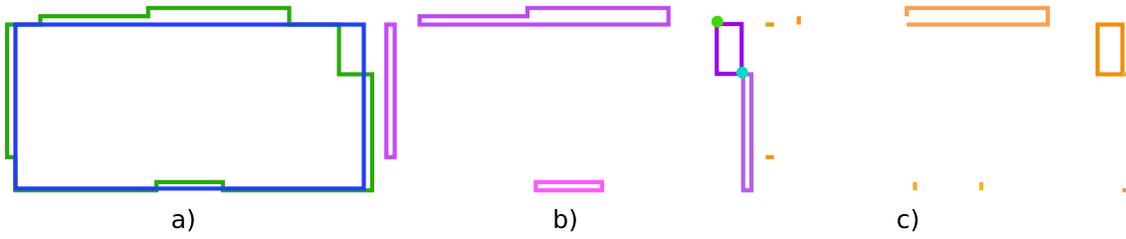


FIGURE 7.3 – La figure a) présente les deux contours à coder de façon différentielle. La figure b) présente les composantes connexes dont le codage correspond à la somme des codages des descriptions différentielles vers l’avant et vers l’arrière sans tolérance par rapport à des imprécisions géométriques. Enfin, la figure c) présente les éléments à coder restants après prise en compte du bruit géométrique. Dans la figure b), le point vert représente un nœud simple, et le point bleu correspond à un point multiple pour lequel il faudra choisir quel chemin prendre pour trouver les cycles qui nous intéressent.

possible la région. Les contours d’erreurs sont ainsi successivement retirés de l’ensemble d’éléments de contour jusqu’à ce qu’aucun cycle n’existe.

7.4 Modèle radiométrique spatio-temporel

La radiométrie des objets des STIS n’est pas invariante dans le temps comme dans la plupart des contenus vidéo. Par ailleurs les évolutions considérées présentent des fréquences trop élevées par rapport aux fréquence d’échantillonnage au sens de Shannon. Nous allons donc supposer dans cette partie que les radiométries des différentes instances d’un objet spatio-temporel sont indépendantes.

Nous considérons par ailleurs une distribution normale des valeurs des pixels sur une région spatiale. La description des valeurs des pixels d’une région consiste donc en la description des paramètres du modèle, et en la description des valeurs des pixels selon ce modèle. On dénomme ces deux termes longueur de description des paramètres, et longueur de description radiométrique. Les longueurs de description associées sont, ¹ :

$$\mathcal{L}_{\text{param}}(r_t) = \frac{d}{2} \left(1 + \frac{d+1}{2} \log(|r_t|) \right) \quad (7.11)$$

et :

$$\mathcal{L}_{\text{radiom}}(r_t) = \frac{|r_t|}{2} (d + \log(2\Pi) + d \log(\Gamma_{r_t})) \quad (7.12)$$

Et la longueur de description totale s’écrit :

$$\mathcal{L} = \sum_{r \in \text{Im}_1, \dots, \text{Im}_T} \mathcal{L}_{\text{geom}}^{\text{ST}}(r) + \mathcal{L}_{\text{param}}(r) + \mathcal{L}_{\text{radiom}}(r) \quad (7.13)$$

7.5 Minimisation de la longueur de description totale

Maintenant que nous disposons de la longueur de description totale, nous devons trouver la séquence de partitions correspondant au minimum de cette longueur. Comme nous l’avons noté au ch.5, la méthode d’optimisation la plus efficace est celle de [Galland et al.,

1. Pour plus de détails, le lecteur est renvoyé aux §5.2.1.2 et 5.2.1.3

2003]. Il s'agit d'une procédure itérative dont une itération se décompose en trois étapes : fusion de régions, déformation des contours, et simplification des contours. L'efficacité de cette méthode d'optimisation est principalement liée au calcul d'une image de somme des statistiques requises de sorte que le calcul de la longueur de description radiométrique de la région se résume à un parcours des contours et non pas des régions. Dans notre cas, la partie la plus coûteuse est le calcul du cylindre de la nouvelle région et de la longueur de description temporellement dépendante qui en découle. Comme nous l'avons précédemment remarqué, les mesures de distances d'appariement entre régions basées région permettent une mise à jour rapide des distances suite à une fusion car ces distances se calculent à partir de l'aire des régions et de leur intersection. Or ces grandeurs sont mises à jour suite à une fusion par une simple addition. En effet, $(r_1^t \cup r_2^t) \cap r_2^{t'} = (r_1^t \cap r_2^{t'}) \cup (r_2^t \cap r_2^{t'})$, or $r_1^t \cup r_2^t = \emptyset$ donc $|(r_1^t \cup r_2^t) \cap r_2^{t'}| = |(r_1^t \cap r_2^{t'})| + |(r_2^t \cap r_2^{t'})|$. En revanche, l'intersection avec deux régions dont le contour a été déformé ne se déduit pas facilement des intersections antérieures. Nous proposons donc une stratégie d'optimisation par fusion de régions.

7.5.1 Présentation de l'algorithme

De façon à fusionner et mettre à jour efficacement les paramètres nécessaires au calcul de la longueur de description, nous proposons l'utilisation d'un SOTAG. Un SOTAG est une structure permettant de représenter les relations temporelles et spatiales entre les objets spatiaux. Nous stockerons dans cette structure les paramètres nécessaires au calcul des coûts de fusion, et effectuerons les fusions dans ce graphe plutôt que dans les images. Les attributs des nœuds sont, comme dans le cas 2D, les statistiques suffisantes jusqu'à l'ordre 2. Dans le cas bidimensionnel, le nombre de segments des frontières entre régions suffisaient pour le calcul des coûts de description géométriques. Dans le cas multi-temporel, en revanche, il faudra stocker, pour chaque arc spatial, le code de Freeman correspondant à la frontière associée afin de calculer les longueurs de description géométriques conditionnelles. Enfin, nous stockerons dans les arcs temporels un bit codant la dépendance temporelle de façon à obtenir les cylindres des régions en parcourant les arcs sortants.

La longueur de description considérée est spatio-temporelle et dépend de plusieurs images. Une simple fusion spatiale de deux régions présentant une bonne mise en correspondance provoquera une augmentation de la longueur de description spatio-temporelle géométrique car on perd la bonne correspondance. Or un schéma d'optimisation par fusion pour minimiser une longueur de description d'image dans le but de segmenter correspond à une simplification progressive du modèle. Ainsi, la longueur de description de l'état initial pris comme sur-partition triviale,² présente une longueur de description élevée du modèle et une description de l'image sachant le modèle nulle, et au fur à mesure de l'algorithme, les fusions simplifient le modèle et augmentent en contrepartie le deuxième terme jusqu'à atteindre le compromis idéal. Il est donc nécessaire que l'opération considérée à chaque itération fasse diminuer la longueur de description globale, or la longueur de description radiométrique croît forcément suite à une fusion. Le fait que certaines fusions provoquent une augmentation de la longueur de description du modèle risque de faire converger l'algorithme vers un minimum local, soit parce que l'optimisation sera terminée trop tôt, soit parce qu'on aura suivi un mauvais chemin d'optimisation.

De façon à régler ce travers, nous devons considérer une série de fusions comme opération élémentaire à effectuer à chaque itération. Nous appelons cette opération élémentaire *Multi-fusion*. En effet, à partir d'une situation stable, en effectuant simultanément plu-

2. une région correspond à un pixel

sieurs fusions dans chaque image de référence, on assure que la longueur de description géométrique ST décroisse. On appelle *fusion clé* la fusion initiale, et *fusions déclanchées* les autres. Les cas d'objets présentant des bonnes mises en correspondances avec toutes les images de la STIS nécessiteront une fusion dans chacune des images en même temps. Il nous faudra donc trouver les fusions déclanchées par une fusion dans toutes les autres images, or comme nous l'avons précédemment remarqué, l'utilisation du graphe intersection complet n'est pas envisageable.

Remarquons par ailleurs qu'une fusion dans une image a non seulement pour effet d'augmenter la longueur de description, mais aussi de diminuer fortement le coût de fusion des régions correspondantes dans les autres images. Nous proposons donc de restreindre temporellement les multi-fusions à une période temporelle d'influence de la fusion-clé que l'on prendra égale à W . Le coût associé à une multi-fusion est lui aussi temporellement restreint et correspond au coût de longueur de description spatio-temporelle d'un sous-graphe correspondant à un découpage spatio-temporel défini par la région clé. On appelle *arc d'ancrage* un arc liant un nœud du sous-graphe à un nœud du graphe privé du sous-graphe. Après multi-fusion dans le sous-graphe, sa réinsertion dans le graphe nécessite une mise à jour des arcs d'ancrage et des longueurs de description de leurs extrémités. Ces mises à jour consisteront pour certains d'entre eux à une forte réduction du coût de description qui aura pour effet de propager implicitement la multi-fusion à travers toute la STIS.

Nous proposons donc un algorithme itératif initialisé par une séquence de sur-partitions. Lors de l'initialisation, on calcule, pour chaque arc spatial, les coûts de la longueur de description du sous-graphe associé. Ces multi-fusions sont ensuite insérées dans une pile de priorité ordonnée par coût croissant. Rappelons que le but étant de faire décroître la longueur de description, plus le coût de multi-fusion est négatif, plus la multi-fusion est intéressante. On dépile ensuite progressivement les multi-fusions tant que le coût associé est négatif en effectuant les multi-fusions correspondantes, puis en mettant à jour les différentes propriétés du graphe, les coûts de multi-fusion et l'ordre dans la pile. Le principe de l'optimisation est présenté figure 7.4.

7.5.2 Estimation des multi-fusions

De façon à minimiser la longueur de description en profitant de la redondance temporelle, nous devons fusionner conjointement les régions de différentes images. En effet, une fusion spatiale dans une image affecte la mise en correspondance avec les régions des W images précédentes qui lui sont temporellement liées. Malgré la restriction du nombre d'images de référence considérées, il peut toujours y avoir un nombre important de régions affectées par une fusion, et donc plusieurs fusions influencées. La multi-fusion associée à une certaine fusion-clé m^* est la combinaison parmi l'ensemble de ces fusions influencées permettant d'obtenir la plus importante décroissance de longueur de description spatio-temporelle. Cette combinaison peut être représentée par un vecteur de vecteur de bits, A , associés à l'ensemble des fusions impliquées dans une multi-fusion, $\mathcal{M}_d(m^*)$: la fusion-clé, ainsi qu'un ensemble de régions influencées. On dit qu'une fusion est *active* si elle appartient à la combinaison optimale. Dans ce cas, le bit correspondant est à 1.

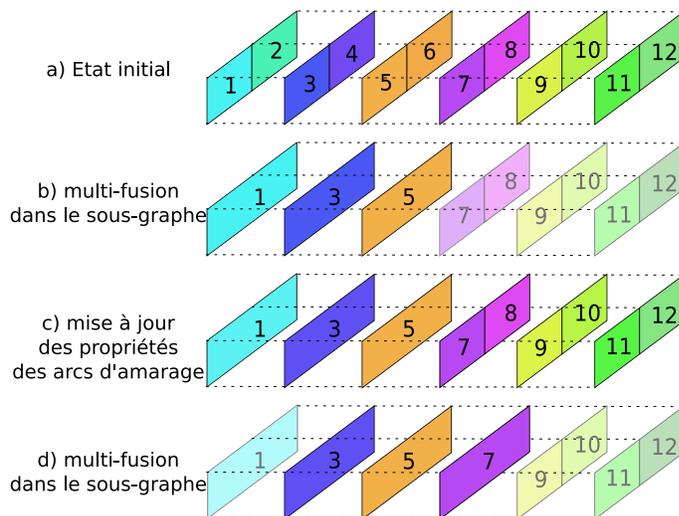


FIGURE 7.4 – Principe de l’optimisation pour $W = 2$: La fusion clé est $m^* = (r_5, r_6)$. Cette fusion influence les deux fusions antérieures $m_1 = (r_3, r_4)$ et $m_2 = (r_1, r_2)$. La multi-fusion effectuée sur le sous-graphe constitué des nœuds $r_1, r_2, r_3, r_4, r_5, r_6$ consiste en ces trois fusions (deuxième ligne). Le sous-graphe ainsi modifié est réinséré dans le graphe, et les coûts de fusion de (r_7, r_8) et (r_9, r_{10}) sont alors mis à jour (ligne 3). Les longueurs de descriptions de r_7 et r_8 sont alors très coûteuses, et en conséquence, leur coût de fusion est très intéressant. La multi-fusion de la fusion-clé (r_7, r_8) –qui n’influence aucune fusion– est alors effectuée. Et ainsi de suite.

Une multi-fusion $M(m^*)$ s’écrit donc $M(m^*) = (\mathcal{M}_d(m^*), A(m^*))$, avec :

$$\mathcal{M}_d(m^*) = \begin{bmatrix} m^* \\ m_{t-1}^1, \dots, \dots, m_{t-1}^{n_1} \\ m_{t-2}^1, \dots, m_{t-2}^{n_2} \\ \vdots \\ m_{t-W}^1, \dots, m_{t-W}^{n_1} \end{bmatrix}$$

$$A(m^*) = \begin{bmatrix} 1 \\ 10110\dots\dots, 101 \\ 1011011 \\ \vdots \\ 101, \dots, 011 \end{bmatrix}$$

7.5.2.1 Fusion Influencée

Dans ce paragraphe, nous décrivons le calcul de l’ensemble des fusions influencées, $\mathcal{M}_d(m^*)$. On dira qu’une fusion m_i est influencée par une fusion m^* si le coût de fusion associé à la longueur de description géométrique temporellement dépendante est plus faible quand m_i et m^* sont effectuées de façon jointe que quand elles sont effectuées de façon indépendantes. L’ensemble des fusions influencées par m^* est noté $\mathcal{M}_d(m^*)$. Formellement :

$$m_i \in \mathcal{M}_d(m^*) \iff \begin{cases} \delta \mathcal{L}_{\text{geom}}^{\text{TD}}(m_i, m^*) < \delta \mathcal{L}_{\text{geom}}^{\text{TD}}(m_i) + \delta \mathcal{L}_{\text{geom}}^{\text{TD}}(m^*) \\ \text{ET} \\ \delta \mathcal{L}_{\text{geom}}^{\text{TD}}(m_i, m^*) < \delta \mathcal{L}_{\text{geom}}^{\text{TD}}(m^*) \end{cases} \quad (7.14)$$

Dans l'exemple de la figure 7.5, les fusions influencées par (r_5, r_6) sont les fusions (r_2, r_4) et (r_3, r_4) , et pas (r_2, r_3) bien que cette fusion présente le coût de fusion jointe le plus faible. Calculer la multi-fusion revient à calculer la meilleure combinaison de fusions parmi toutes ces fusions influencées. Le nombre de combinaisons possibles est $\sum_{t \in \llbracket t-1, t-W \rrbracket} 2^{N_d^{t'}}$ où $N_d^{t'}$ est le nombre de fusions influencées dans l'image t' . De façon à réduire encore le nombre de combinaisons, nous considérons une multi-fusion constituée d'au plus une fusion par temps. Le nombre de combinaisons descend alors à 2^W . On a alors :

$$\mathcal{M}_d(m^*) = \begin{bmatrix} m^* \\ m_{t-1} \\ m_{t-2} \\ \vdots \\ m_{t-W} \end{bmatrix} \quad (7.15)$$

$$A(m^*) = \begin{bmatrix} 1 \\ 0 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \quad (7.16)$$

Propriété 7.5.1 : *Une fusion dont l'une des régions a une intersection vide avec $\bigcup_{r \in m^*} r$ n'est pas une fusion influencée*

En effet, si l'une des régions a une intersection vide avec la région fusionnée de m^* , alors, la deuxième condition de 7.17 n'est plus vérifiée.

Propriété 7.5.2 : *Soit r_1^* et r_2^* les deux régions de la fusion m^* . Une fusion dont aucune des deux régions ni la réunion n'est meilleure correspondante ni de r_1^* , ni de r_2^* , ni de $r_1^* \cup r_2^*$ n'est pas une fusion influencée*

En effet, le coût de longueur de description géométrique temporellement dépendant dépend de la longueur de description géométrique temporellement dépendante des régions séparées et des régions fusionnées. Et ces longueurs de description géométrique dépendent du cylindre, c'est-à-dire des meilleures correspondantes.

Propriété 7.5.3 : *Deux régions qui ne sont pas spatialement voisines ne peuvent pas fusionner.*

Ces trois propositions permettent de définir l'ensemble des paires de régions possiblement influencées. On note cet ensemble $Cand(m^*)$.

La sélection d'une unique fusion influencée par l'image de référence est effectuée en minimisant la différence des coûts de fusion jointe et disjointe de la longueur de description géométrique temporellement dépendante :

$$\begin{array}{l} \hat{m}_i \text{ est la fusion} \\ \text{la plus influencée} \\ \text{par } m^* \end{array} \iff \hat{m} = \underset{\substack{m_i \in Cand(m^*) \\ m_i \in \mathcal{M}_d(m^*)}}{\operatorname{argmin}} \delta \mathcal{L}_{\text{geom}}^{\text{TD}}(m_i, m^*) \quad (7.17)$$

Remarquons que dans certains cas, il n'y aura pas de fusion déclanchée. Ce sera le cas par exemple si les régions de m^* sont incluses dans une même région dans l'image de référence.

Une fois la séquence des fusions influencées obtenue, nous devons calculer la longueur de description spatio-temporelle du sous-graphe associé pour chacune des combinaisons possibles.

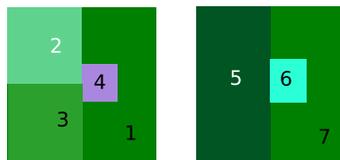


FIGURE 7.5 – Les fusions influencées par la fusion (r_5, r_6) sont les fusions (r_2, r_4) et (r_3, r_4) . La fusion sélectionnée est (r_3, r_4) .



FIGURE 7.6 – Exemple de fusion inactive : la fusion influencée par la fusion de r_3 et r_4 est la fusion de r_1 et r_2 , cependant, cette fusion coûtera cher en description géométrique temporellement dépendante et sera donc inactive dans la multi-fusion déclenchée par (r_1, r_2) .

7.5.2.2 Sous-graphe associé à une multi-fusion

Les nœuds influencés indirectement par une fusion correspondent aux nœuds dont la longueur de description géométrique temporellement dépendante dépend des régions impliquées dans les fusions actives. Le sous-graphe associé à une multi-fusion est le sous-graphe induit par les nœuds des temps $t - k$ où $k \in \llbracket 0, W \rrbracket$ influencés de façon directe ou indirecte par la fusion clé. Lors de l'estimation de l'activité des fusions d'une multi-fusion, on considère toutes les fusions influencées comme actives pour le calcul du sous-graphe de façon à obtenir des coûts de description comparables.

Le sous-graphe est donc le sous-graphe induit par :

- Les régions des fusions actives
- Les voisines temporelles passées des régions des fusions actives

7.5.2.3 Estimation de l'activité des fusions de la multi-fusion

Dans ce paragraphe, nous décrivons le calcul du vecteur de bits, $A(m^*)$, associé à une multi-fusion. Les fusions influencées sont déterminées sur un critère géométrique temporellement dépendant uniquement. Il reste donc à choisir quelle sélection parmi ces fusions influencées permet de décroître le plus la longueur de description totale du sous-graphe associé. La figure 7.6 présente une fusion influencée qui sera désactivée à cause de la mauvaise correspondance qu'elle entraîne.

Afin de déterminer la configuration optimale, on calcule, pour toutes les configurations possibles du vecteur de bits associé à la multi-fusion le coût correspondant pour la longueur de description dans le sous-graphe. Les coûts radiométriques, paramétriques et géométriques temporellement localisés sont les mêmes que dans le cas d'une seule image et sont détaillés §5.2.2. Ces coûts sont calculés directement sans passer par les longueurs de description correspondantes avant et après fusion car on peut les formuler analytiquement. Les longueurs de description géométrique temporellement dépendantes dépendent des régions temporellement liées et impliquent une minimisation des distances de forme entre ces régions suivie d'un calcul de coût de transport du contour de la région dont la

forme est la plus similaire au contour de la région courante. Une formulation analytique de ce coût n'est pas triviale et supposerait de stocker beaucoup de variables intermédiaires, aussi, nous proposons de calculer ce coût par la différence entre la longueur de description temporellement dépendante avant et après fusion.

On calcule donc dans un premier temps la longueur de description temporellement dépendante des régions du sous-graphe, puis on effectue les fusions simultanément à la mise à jour de propriétés nécessaires au calcul des distances de forme et du coût de description conditionnelle des régions. Notons que la longueur de description géométrique temporellement dépendante des nœuds du sous-graphe n'est pas nécessairement la même que celle des mêmes nœuds dans le graphe car le nombre d'images de référence n'est pas le même pour tous les nœuds.

Les propriétés nécessaires au calcul des distances de forme et du coût de description conditionnelle des régions sont les tailles, les contours, et les intersections. Comme nous l'avons précédemment remarqué, les tailles et les intersections sont mises à jour par une simple addition des tailles et intersections correspondantes pour les régions antérieures, et les contours sont mis à jour par une simple concaténation des contours des régions antérieures. Le cylindre, ainsi que la longueur de description géométrique conditionnelle doivent ensuite être recalculés pour les régions résultant d'une fusion. Enfin, pour les régions futures qui leur sont temporellement liées, un recalcul du cylindre doit être effectué, suivi d'un recalcul de la longueur de description conditionnelle si le cylindre a changé, et si une région fusionnée y appartient.

Une fois ces mises à jour effectuées, la longueur de description géométrique conditionnelle du sous-graphe est recalculée, et le coût correspondant est obtenu par différence avec la longueur de description géométrique conditionnelle antérieure.

7.5.3 Multi-Fusion et mise à jour de la pile de priorité

Une fois que la pile de priorité des multi-fusions est calculée, il nous faut la dépiler en effectuant les multi-fusions constituées des fusions actives. Pour ce faire, on commence, comme pour l'estimation de l'activité d'une multi-fusion, par effectuer les fusions actives et mettre à jour : les tailles, et paramètres radiométriques des régions impliquées, ainsi que les arcs temporels et intersections correspondantes, et les arcs spatiaux, et portions de contour correspondantes. On met ensuite à jour :

- les tailles des régions impliquées dans une fusion
- les paramètres radiométriques des régions impliquées dans une fusion
- les arcs temporels dont une extrémité est une région impliquée dans une fusion et les intersections correspondantes
- les cylindres des régions impliquées dans une fusion
- les longueurs de description géométrique dépendantes pour les régions dont le cylindre a été modifié
- les coûts de fusion impliquant une région spatialement voisine d'une région résultante d'une fusion
- les coûts de fusion impliquant une région dont la longueur de description géométrique dépendante a été modifiée

Cette mise à jour a pour effet de placer en tête de la pile les fusions correspondantes dans les images postérieures à celles de la fusion-clé de la multi-fusion ainsi que dans les W images antérieures à la dernière image influencée par la région clé et de propager ainsi la multi-fusion le long de la séquence.

7.6 Conclusion

Cette méthode permet donc de segmenter conjointement les images d'une STIS par codage d'un SOTAG. Il en résulte une séquence de partitions en objets spatiaux présentant une homogénéité spatiale radiométrique et une redondance temporelle de forme. Les séquences de partitions obtenues présentent donc une finesse variable liée à la stabilité temporelle des objets. Les objets ainsi obtenus se rapprochent donc plus des objets sémantiquement pertinents. Un point négatif de la méthode est cependant sa complexité qui rend les calculs extrêmement longs et limite l'applicabilité à grande échelle. Le point critique de l'algorithme en termes de temps de calcul est le calcul des distances de Hausdorff entre les différentes chaînes de contour. On peut donc espérer un gain de temps sensible en améliorant la façon dont sont calculées ces distances ou encore en considérant une autre distance adaptée et plus simple à calculer.

Trouver conjointement les objets et leurs correspondances est un problème très difficile, et nous avons dû effectuer une importante approximation sur le SOTAG en ne considérant que des dépendances 1-1 entre les régions de différents instants. Les objets spatiaux étant maintenant extraits, nous allons pouvoir relaxer cette approximation et rajouter les arcs temporels correspondant à des changements topologiques.

Chapitre 8

Analyse géométrique du graphe d'adjacence temporelle

Dans cette section nous cherchons à mettre en correspondance les différents objets de la séquence de façon transitive. Nous avons vu au §A.2.5 différentes méthodes de mises en correspondance de graphe. Cependant, comme nous l'avons alors souligné, le problème que nous considérons ici est différent du problème de la mise en correspondance d'images stéréo d'une part car aucune contrainte épipolaire n'est ici applicable (du moins pour mettre en évidence les changements des objets), et d'autre part car un isomorphisme de sous-graphe ne permet pas une description complète de l'*histoire* de tous les objets. En effet, le problème que nous considérons ici est une mise en correspondance individuelle de toutes les régions. Chaque région est liée à au moins une région de l'image précédente et de l'image suivante.

Nous avons, dans le chapitre 7, présenté une méthode d'extraction des objets du SOTAG. Cependant, pour ce faire, nous avons eu recours à des approximations fortes concernant les dépendances temporelles. Ces objets étant maintenant déterminés, nous pouvons nous permettre de relâcher la contrainte portant sur le nombre de dépendances temporelles afin d'obtenir un SOTAG se rapprochant plus du SOTAG idéal de la scène.

Les objets sont certes fixés, mais les erreurs de recalage géométrique et de segmentation demeurent, le critère d'intersection n'est donc toujours pas suffisant pour mettre en correspondance les régions. Comme nous l'avons mentionné au §7.2.4, trouver ces correspondances est un problème complexe d'optimisation. En effet, la décision concernant la mise en correspondance 1-1 de deux régions ne dépend pas simplement de la valeur de leur distance de forme mais aussi du contexte. Nous détaillons ce point dans le §8.1. Ce contexte correspond à la composante connexe du SOTAG contenant les régions considérées. Dans le cas où le critère d'intersection est considéré pour l'adjacence temporelle, cette composante connexe peut s'étendre spatialement à toute l'image, ce qui explique la complexité de l'optimisation. Il existe néanmoins certains cas où la distance de forme entre régions est tellement faible qu'on n'a pas besoin du contexte pour prendre la décision de mise en correspondance. Nous proposons de traiter ces cas en premier de façon à créer des composantes connexes de taille réduite déconnectées du reste du graphe ce qui a pour conséquence de réduire la taille des composantes connexes des autres nœuds du graphe et donc de simplifier la prise de décision.

L'algorithme que nous proposons part du SOTAG dont les arcs correspondent à une intersection non vide, puis on élague progressivement les arcs en prenant les décisions les plus simples en premier par simple seuillage d'une distance de forme de sorte que les cas

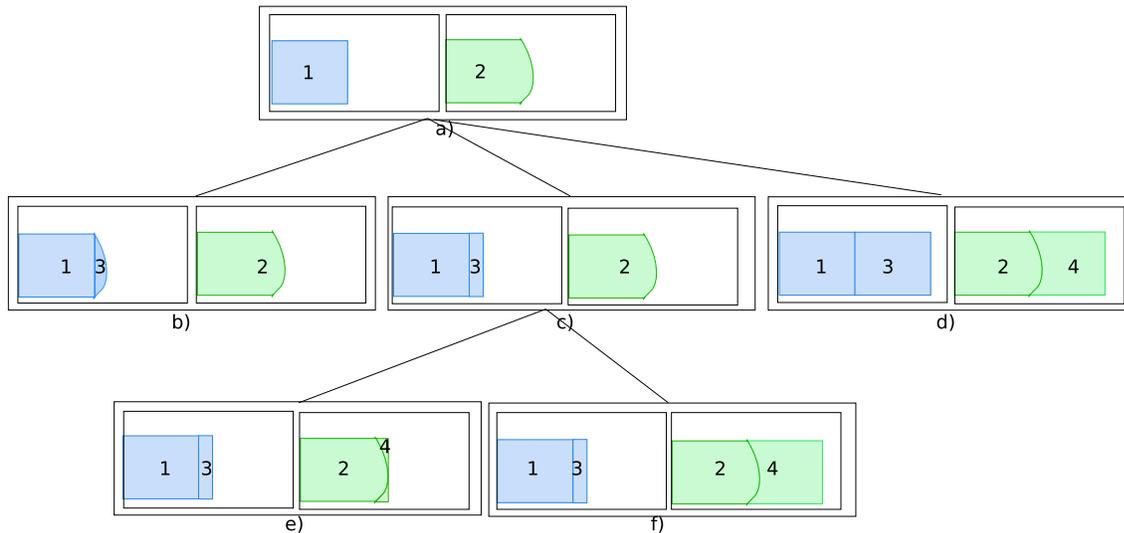


FIGURE 8.1 – La première ligne présente deux régions dont la mise en correspondance est ambiguë. La deuxième ligne présente différents voisinages possibles pour la région r_1 . Le cas c) n'est plus ambigu : r_1 et r_3 se sont fusionnées en r_2 . De même, le cas d) n'est pas ambigu : il s'agit d'une combinaison de fusions entre $\{r_1, r_2\}$ et $\{r_3, r_4\}$. Le cas b), en revanche, est ambigu. La troisième ligne présente deux contextes différents possibles pour la région r_2 . Le cas e) correspond au cas d), et le cas f) sera décrit par une fusion, comme le cas b), mais avec une erreur géométrique.

ambigus traités à la fin appartiennent à une composante connexe de taille réduite.

8.1 Importance du contexte

En raison des erreurs de segmentation ou de recalage, la mise en correspondance de deux régions, d'une région et d'un groupe de régions, ou de deux groupes de régions, dépend du voisinage spatial respectivement des régions, et groupes de régions. En effet, les différences de forme entraînent la superposition partielle des régions voisines de la région ou du groupe de régions considérées. La figure 8.1 illustre l'importance du contexte.

Nous voyons donc l'importance du contexte dans la prise de décision, et c'est ce qui contribue à rendre le problème difficile.

8.2 Algorithme

Comme nous l'avons précédemment remarqué, la recherche d'une correspondance n-m est difficile. Cependant, il existe une partition des nœuds et des arcs du SOTAG en nœuds et arcs impliqués respectivement dans une correspondance 1-1, 1-n, et n-m. Ainsi, en commençant par traiter les correspondances 1-1, et 1-n, les nœuds et arcs restant non attribués à un type de correspondance seront ceux qui sont impliqués dans une correspondance n-m. Nous proposons donc dans un premier temps de déterminer et d'isoler toutes les correspondances 1-1 et 1-n de façon à en déduire par la suite les correspondances n-m sans avoir à les calculer.

Par ailleurs, nous avons vu que la prise de décision de mise en correspondance dépend

du contexte. Dans certains cas, cependant, la décision ne nécessite pas de regarder le voisinage spatial. C'est le cas par exemple d'une correspondance exacte (distance de forme nulle). Nous proposons donc un algorithme de type *winner-take-all* pour mettre en évidence d'abord les correspondances 1-1, puis les correspondances 1-n. Cependant, comme nous l'avons vu au §A.2.5, cette approche présente l'inconvénient d'obtenir des gagnants par défaut en fin de parcours. Afin de réduire cet effet non désirable, nous proposons une approche itérative selon laquelle les correspondances choisies sont non seulement les meilleures, mais en plus, ont une mesure de correspondance inférieure à un seuil que l'on fait croître entre deux itérations. A chaque itération it , on valide les arcs permettant des mises en correspondance non ambiguës, c'est-à-dire dont les distances de forme sont en dessous d'un seuil s_{it} et on invalide les arcs de façon à les isoler de leur contexte. Lors de l'itération suivante, on valide les arcs restants non validés dont la mise en correspondance est inférieure à un seuil s_{it+1} tel que $s_{it+1} > s_{it}$. De plus, une mise en correspondance 1-1 requiert un plus petit voisinage qu'une correspondance 1-n. A chaque itération, nous allons donc traiter dans un premier temps les correspondances 1-1, et dans un second temps les correspondances 1-n.

8.2.1 Correspondances 1-1

Remarquons que la relation de correspondance 1-1 est symétrique : $(r_1 \leftrightarrow r_2) \iff (r_2 \leftrightarrow r_1)$. Nous n'allons donc traiter ces correspondances que dans un sens. Nous avons choisi le sens vers le passé.

Nous avons vu que l'algorithme est itératif, de sorte qu'à chaque itération des nœuds sont attribués à un type de mise en correspondance, des arcs sont validés, et d'autres invalidés. Soit \mathcal{U}^t l'ensemble des nœuds de temps t , et \mathcal{U} l'ensemble des nœuds. Notons $\mathcal{U}_d^+ \subset \bigcup_{t=1}^{T-1} \mathcal{U}^t$ (réciproquement $\mathcal{U}_d^- \subset \bigcup_{t=0}^{T-2} \mathcal{U}^t$) l'ensemble des nœuds dont la correspondance vers le passé (réciproquement vers le futur) n'est pas déterminée, et \mathcal{U}_p^+ (réciproquement \mathcal{U}_p^-) l'ensemble des nœuds traités vers le passé (réciproquement vers le futur) ($\mathcal{U} = \mathcal{U}_d \cup \mathcal{U}_p$). De même, notons \mathcal{H}_d l'ensemble des arcs non validés, et \mathcal{H}_p l'ensemble des arcs validés.

Cette étape consiste à parcourir tous les nœuds u_0^t de \mathcal{U}_d^- de temps t non nul, et à calculer la distance de forme entre la région correspondant à ce nœud et sa meilleure correspondance u_0^{t-1} dans l'image précédente liées par un arc de \mathcal{H}_d . Si cette distance est en-dessous d'un certain seuil s_{it} , alors les arcs et nœuds impliqués dans la correspondance 1-1 sont déplacés vers les ensembles respectifs de nœuds et arcs traités. Maintenant que la décision de correspondance 1-1 est prise pour cette région, on procède à la suppression des arcs liés au bruit géométrique. De plus, afin de tirer parti du contexte, si le nœud courant n'a plus qu'une correspondance candidate, et réciproquement que le nœud de cette correspondance n'a pour correspondance candidate que le nœud courant, ces nœuds sont mis en correspondance sans calculer leur distance.

- l'arc de la correspondance 1-1 est supprimé de \mathcal{H}_d et ajouté à \mathcal{H}_p
- les arcs sortant de u_0^{t-1} vers le futur dont l'extrémité est différente de u_0^t sont supprimés
- les arcs sortant de u_0^t vers le passé dont l'extrémité est différente de u_0^{t-1} sont supprimés
- u_0^t est supprimé de \mathcal{U}_d^- et inséré dans \mathcal{U}_p^-

8.2.2 Correspondances 1-n

Notons r^t une région de l'instant t . Dans le cas où il n'y a pas d'erreur de localisation des contours, dire que la région r^t résulte d'une division équivaut à $\exists r^{t-1} | r^t \subset r^{t-1}$. Et de même, dire que la région r^{t-1} est impliquée dans une fusion équivaut à $\exists r^t | r^{t-1} \subset r^t$.

Nous proposons d'étendre ces concepts d'inclusion et de contenance aux cas où il y a des erreurs de localisation. On définit ainsi la relation \subset_t :

$$r_1 \subset_t r_2 \iff \frac{|r_1 \cap r_2|}{|r_1|} > t \quad (8.1)$$

Traiter les correspondances 1-n signifie, pour tous les nœuds u_0^t de \mathcal{U}_d^- , déterminer si la région correspondant à u_0^t est incluse ou contient la région associée à l'un de ses prédécesseurs. Si tel est le cas, il nous faut déterminer le groupe de régions impliquées dans la fusion ou résultant d'une division. Le cas de la fusion est le même que celui de la division, mais avec une autre convention de sens pour le temps. Nous allons donc présenter simplement comment obtenir le groupe de régions de $t - 1$ impliquées dans une fusion résultant en la région correspondant à u_0^t .

Les prédécesseurs de u_0^t liés par des arcs de \mathcal{H}_d sont d'abord classés par distance logarithmique (définie en 7.3.1) décroissante. Ils sont ensuite successivement agglomérés tant que la distance entre la région résultant de la réunion du groupe de régions et la région de u_0^t est inférieure à s . Une fois ce seuil atteint, on continue l'agglomération tant que la distance décroît.

Si un groupe de régions dont la forme est distante de moins de s de la région de u_0^t est ainsi trouvé :

- l'ensemble des arcs de u_0^t vers les nœuds correspondant aux régions du groupe est déplacé de \mathcal{H}_d vers \mathcal{H}_p
- les arcs vers le passé partant de u_0^t dont l'extrémité n'est pas un nœud du groupe sont supprimés
- les arcs vers le futur partant d'un nœud du groupe dont l'extrémité n'est pas u_0^t sont supprimés
- u_0^t est déplacé de \mathcal{U}_d^- vers \mathcal{U}_p^- .

8.2.3 Post-traitement

Le SOTAG tel que nous l'avons construit peut présenter des nœuds isolés. Afin de fournir une description complète, il nous faut relier ces nœuds. Nous proposons donc un post-traitement consistant simplement à lier ces nœuds à la région présentant la plus faible distance de forme. Notons qu'il peut néanmoins être intéressant de garder la description des événements correspondant au SOTAG avant pré-traitement pour les nœuds qui n'étaient pas isolés, par exemple pour le cas de la figure 8.2,

La méthode proposée utilise des seuils pour les mesures de correspondance. En conséquence, il nous faut une mesure qui permette de comparer des correspondances n'ayant aucun nœud en commun : $(n_{t,i}, n_{t+1,j})$ avec $(n_{t,k}, n_{t+1,l})$. La distance de Hausdorff est donc un bon choix de mesure de correspondance.

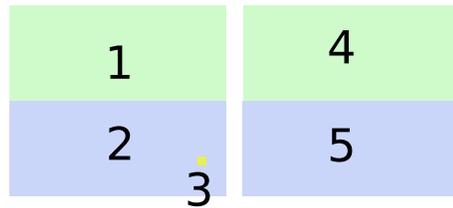


FIGURE 8.2 – Le SOTAG avant pré-traitement considère deux correspondances 1-1 : $r_1 \leftrightarrow r_4$ et $r_2 \leftrightarrow r_5$. Le nœud correspondant à la région r_3 est donc isolé. Le SOTAG après post-traitement reliera le nœud de r_3 au nœud de r_5 modifiant ainsi l'événement spatial de r_2 en l'impliquant dans une fusion.

Chapitre 9

Recherche de séquences de caractéristiques fréquentes dans le graphe d'adjacence spatio-temporelle

Lors de la segmentation jointe des images, nous avons exploité la redondance de forme, et supposé l'indépendance des différentes instances des objets. L'hypothèse d'indépendance alors formulée est liée à un manque d'information sur le type d'évolution et sur la calibration. Cependant, les différentes instances des caractéristiques d'un objet sont liées, et nous proposons ici d'analyser ces évolutions radiométriques des SO maintenant définis et temporellement liés. Nous avons vu au ch.4 que les caractéristiques des STO changent selon des processus et des événements. L'analyse des caractéristiques revient donc à estimer les processus et les événements des STO.

Pour analyser ces évolutions, nous avons choisi une méthode de recherche de motifs fréquents dans les séquences de caractéristiques exprimées par les STO au cours du temps. Ce choix d'analyse se justifie par le manque d'information disponible dans les données. Nous expliquons ce choix ci-après en présentant différentes options pour l'analyse, ainsi que d'autres approches adoptées dans l'état de l'art que nous avons rejetées en raison des contraintes des données.

Nous avons vu au §A.2.4.1 une approche permettant d'analyser les évolutions temporelles de caractéristiques radiométriques utilisant des modèles d'évolution. Soit les données sont filtrées, de façon à permettre l'extraction de paramètres phénologiques permettant l'identification de l'évolution. Soit le modèle est directement ajusté sur les données. Cette approche suppose donc un *a priori* sur les données. La scène que nous considérons est hétérogène, de telles approches ne sont donc pas applicables.

D'autres approches consistent à classifier les évolutions temporelles de façon non supervisées, comme dans [Robin et al., 2008, Bouzidi et al., 1998]. Mais la fréquence d'échantillonnage de la séquence ADAM est trop faible pour ces approches, d'ailleurs, dans ces travaux, une séquence à fréquence temporelle plus élevée (NOAA) était fusionnée à la séquence ADAM pour permettre son analyse temporelle.

L'approche duale à l'identification des processus est la détection des événements. L'information concernant tous les événements n'est cependant pas contenue dans les données. Ainsi, un événement du type *semis* n'est pas présent dans les données car on n'observe que sa conséquence : la levée du semis qui a lieu plus tard. Il est donc impossible d'identifier tous les événements sans information *a priori*. Les données sans autre connaissance

permettent donc au mieux d'identifier des évolutions constituées de k processus et $k - 1$ événements avec $k \geq 1$. Il s'agit donc d'un objectif théoriquement possible. Mais qu'en est-il de la pratique ?

A défaut de les détecter tous, on peut chercher à détecter les événements visibles comme étant des changements radiométriques rares. On ne dispose pas d'assez d'échantillons temporels pour suivre et surtout interpoler l'évolution d'un pixel dans le temps, mais on peut néanmoins détecter des anormalités, mises en évidence par un comportement statistiquement cohérent d'un ensemble de pixels. Le manque d'information temporelle est ainsi compensé par la redondance intra-classe : on ne peut pas interpoler l'évolution des pixels dans le temps, en revanche, on peut détecter des évolutions statistiquement rares, et des tendances statistiquement fréquentes. Cette approche, basée sur des transitions, est cependant très sensible aux apparitions intempestives de valeurs divergentes dont les sources sont multiples dans notre méthode : l'apparition de nuages non conformes à l'hypothèse de scène d'arrière-plan, des décalages temporels entre évolutions, des erreurs de segmentation et enfin des erreurs d'appariement.

On pourrait donc essayer d'identifier ces séquences de processus par détection de changements radiométriques (cf. §A.2.4.4). Cette approche dépend cependant d'un réglage du seuil de détection qui peut s'avérer difficile. En effet, les données ne sont pas intercalibrées par rapport aux STO d'intérêt (variation atmosphérique et d'acquisition). Et une calibration des images n'est pas évidente car on ne peut supposer d'invariance globale des radiométries des images en raison des évolutions radiométriques des objets qui la composent ce qui rend inapplicables les méthodes d'intercalibration classiques. Une approche de détection de changement par seuillage d'images de différence est donc aussi inadaptée.

[Héas, 2005] a proposé une méthode de classification d'évolutions dans l'espace MT réduit. L'approche est cependant pixellaire, et bien que les structures trouvées soient généralement compactes, elles présentent un bruit *poivre-et-sel* lié à cette approche. Dans [Gueguen, Le Men and Datcu, 2006], nous avons proposé une méthode de classification de textures pour des séquences de SO. La longueur des évolutions était cependant fixée, ce qui est un facteur limitant les types de structures descriptibles. Enfin, ces méthodes sont adaptées à des espaces vectoriels et la structure de graphe a tendance à augmenter artificiellement la similarité intra-classe de sorte que des structures peu fréquentes peuvent être comptabilisées comme fréquentes. De plus, le cas d'évolutions se succédant n'est pas pris en compte lorsqu'on impose la longueur de la séquence.

Dans [Julea et al., 2006, 2008, Le Men et al., 2008], une analyse temporelle permettant de prendre en compte des successions d'évolutions et des supports temporels non compacts ont été proposées. Après discrétisation de l'espace des valeurs, les séquences de symboles correspondant à chaque pixel forment une base de séquences dans lequel un algorithme de recherche de séquences fréquentes sous contraintes ([Zaki, 2000a]) est appliqué. Les évolutions fréquentes permettent de caractériser un ensemble de pixels à des positions spatiales et temporelles particulières. Afin de prendre en compte la structure spatiale des objets de la STIS, nous avons intégré cette information dans [Le Men et al., 2008] aux objets spatiaux issus d'une segmentation des images. Cependant, dans ce cas, aucune cohérence temporelle géométrique entre les SO n'était considérée : les SO des différentes images étaient classifiés par rapport aux évolutions des pixels qu'ils contiennent à un instant, indépendamment des changements de forme qu'ils peuvent subir et qui ont pour effet de modifier l'influence de certains pixels. Nous proposons donc une extension de ce type de méthodes aux SOTAG en considérant comme base de séquences non plus les séquences des valeurs discrétisées de pixels au cours du temps, mais les séquences des valeurs discrétisées

des caractéristiques des STO correspondants aux chemins du SOTAG.

9.1 Classification des caractéristiques TL des régions

Afin de mettre en évidence des évolutions fréquentes, il nous faut dans un premier temps discrétiser l'espace des caractéristiques TL. Dans [Le Men et al., 2008, Julea et al., 2008], la discrétisation est effectuée indépendamment pour chaque canal, tandis qu'elle est effectuée globalement pour la séquence dans [Julea et al., 2006]. La discrétisation globale sur la séquence suppose que les valeurs aux différents instants sont comparables, c'est-à-dire que les images sont inter-calibrées. Etant donné que l'inter-calibration est fortement perturbée par les évolutions locales des STO, nous avons supposé ici que les valeurs des différentes images ne sont pas comparables, et nous avons effectué les discrétisations image par image indépendamment.

Discrétiser indépendamment chaque canal, et supposer par la suite des motifs TL de trois symboles (correspondant aux trois canaux) comme c'est le cas dans [Julea et al., 2006, 2008] n'est d'une part pas optimal en terme de représentation de l'espace multi-canal et d'autre part complique la structure des motifs à rechercher. Nous avons donc choisi une discrétisation de l'espace multi-canal directe et non pas canal par canal. En contrepartie, ce faisant, on perd le lien entre l'ordre des valeurs initiales et l'ordre des valeurs discrétisées.

Les caractéristiques d'un objet spatial modélisent les valeurs des pixels de la région associée. Regrouper deux objets spatiaux revient donc à considérer que leurs pixels suivent le même modèle.

Nous modélisons les données par un mélange de gaussiennes. Le modèle de mélange de gaussiennes suppose l'indépendance des pixels. On a alors, pour un ensemble de pixels X , des vecteurs moyens μ_k et des matrices de covariance Γ_k : $p(X|\mu_1, \dots, \mu_k, \Gamma_1, \dots, \Gamma_k) = \prod_{x \in X} \sum_{k=1}^K p(c_k)p(x|\mu_k, \Gamma_k, c_k)$. c_k est une variable latente représentant les différentes composantes du modèle. Cependant, notre objectif étant de regrouper des régions, et pas simplement des pixels, nous imposons de plus à tous les pixels d'une même région, R_i d'avoir le même comportement. On a alors :

$$P(X|\mu_1, \dots, \mu_k, \Gamma_1, \dots, \Gamma_k) = \prod_{j=1}^{|\mathcal{P}|} \sum_{k=1}^K p(c_k) \prod_{i=1}^{|R_j|} p(x_{ij}|\mu_k, \Gamma_k, c_k) \quad (9.1)$$

Les $p(c_k)$ représentent dans ce cas la probabilité qu'une région appartienne à une composante c_k . Cette grandeur étant aussi inconnue, il est classique d'utiliser un algorithme de maximisation de l'espérance (EM) pour estimer les $p(c_k)$, μ_k et Γ_k . La version que l'on propose est un peu différente d'un classique mélange de gaussiennes à cause de l'indépendance conditionnelle des pixels d'une région. Nous explicitons donc ci-dessous les différentes étapes. L'algorithme EM est un algorithme itératif dont la première étape est une initialisation des paramètres des distributions associées aux composantes : $(\mu_k, \Gamma_k, p(c_k))$, $\forall k \in \llbracket 1, K \rrbracket$. On peut alors calculer à partir de ces paramètres l'espérance des données manquantes : $\forall (i, k) \in \llbracket 1, |\mathcal{P}| \rrbracket \times \llbracket 1, K \rrbracket$:

$$(E) : \tilde{P}^{it}(c_k|R_i, p^{it-1}(c_k), \mu_k^{it-1}, \Gamma_k^{it-1}) = \frac{\prod_{j=1}^{|R_i|} p^{it-1}(c_k)p(x_{ij}|\mu_k^{it-1}, \Gamma_k^{it-1})}{\sum_{i=1}^{|\mathcal{P}|} \prod_{j=1}^{|R_i|} p^{it-1}(c_k)p(x_{ij}|\mu_k^{it-1}, \Gamma_k^{it-1})} \quad (9.2)$$

Ceci correspond à l'étape (E). On peut ensuite minimiser la log vraisemblance en annulant ses dérivées partielles par rapport à chacun des paramètres $(\mu_k, \Gamma_k, p(c_k))$. Cette étape,

correspondant à l'étape (M), permet d'obtenir une nouvelle estimation des paramètres utilisant les probabilités calculées à l'étape (E) :

$$\begin{aligned}
 \mu_k^{it} &= \frac{\sum_{i=1}^{|\mathcal{P}|} \tilde{P}^{it}(c_k | R_i, p^{it-1}(c_k), \mu_k^{it-1}, \Gamma_k^{it-1}) \sum_{j=1}^{|R_i|} x_{ij}}{\sum_{i=1}^{|\mathcal{P}|} \tilde{P}^{it}(c_k | R_i, p^{it-1}(c_k), \mu_k^{it-1}, \Gamma_k^{it-1}) |R_i|} \\
 \Gamma_k^{it} &= \frac{\sum_{i=1}^{|\mathcal{P}|} \tilde{P}^{it}(c_k | R_i, p^{it-1}(c_k), \mu_k^{it-1}, \Gamma_k^{it-1}) \sum_{j=1}^{|R_i|} (x_{ij} - \mu_k^{it})^T (x_{ij} - \mu_k^{it})}{\sum_{i=1}^{|\mathcal{P}|} \tilde{P}^{it}(c_k | R_i, p^{it-1}(c_k), \mu_k^{it-1}, \Gamma_k^{it-1}) |R_i|} \\
 p^{it}(c_k) &= \frac{\sum_{i=1}^{|\mathcal{P}|} \tilde{P}^{it}(c_k | R_i, p^{it-1}(c_k), \mu_k^{it-1}, \Gamma_k^{it-1}) |R_i|}{N_x * N_y}
 \end{aligned} \tag{9.3}$$

Ces deux dernières étapes sont itérées jusqu'à convergence. Le nombre optimal de composantes (K) par image est obtenu par un critère MDL [Rissanen, 1978, 1989] :

$$\begin{aligned}
 K &= \underset{P}{\operatorname{argmin}} (\mathcal{L}(\mathbf{X} | \mathbf{\Gamma}, \boldsymbol{\mu}) + \mathcal{L}(\mathbf{\Gamma}, \boldsymbol{\mu})) \\
 \text{avec } \mathcal{L}(\mathbf{\Gamma}, \boldsymbol{\mu}) &= \frac{P \cdot d}{2} \left(1 + \frac{d+1}{2} \right) \log(N_x * N_y)
 \end{aligned} \tag{9.4}$$

Afin de prendre une décision permettant d'affecter une région à une seule composante, on maximise, pour chaque région, R_i $P(c_k | p^{it-1}(c_k), \mu_k^{it-1}, \Gamma_k^{it-1})$:

$$c_{k_i} = \underset{\{c_k\}_{k \in [1, K]}}{\operatorname{argmax}} (P(c_k | p^{it-1}(c_k), \mu_k^{it-1}, \Gamma_k^{it-1})) \tag{9.5}$$

On ne peut définir d'ordre (total ou partiel) sur l'ensemble des valeurs discrétisées liées à leurs valeurs. Nous les identifions donc par un ensemble de symboles que nous ordonnons selon un ordre arbitraire. Etant donné que d'une image à l'autre, les symboles associés aux classes TL sont disjoints, aucune répétition n'est possible dans une séquence temporelle. Ceci nous permet de simplifier grandement l'espace de recherche en ne cherchant non plus des motifs séquentiels fréquents, comme c'était le cas dans [Julea et al., 2006, 2008, Le Men et al., 2008], mais des motifs ensemblistes fréquents, pour lesquels l'ordre n'a pas d'importance. On passe ainsi d'un nombre de motifs possibles de $\sum_{t=0}^{T-1} (N_c)^t$ à $\sum_{t=0}^{T-1} \binom{N_c}{t} = 2^{T-1}$, avec N_c , le nombre de classes TL sur la STIS. ($N_c = \sum_{t=0}^{T-1} N_{ct}$, N_{ct} : nombre de classes TL pour l'image t .)

Ces symboles sont ensuite utilisés pour étiqueter le graphe. On affecte à chaque nœud du graphe le symbole de la classe à laquelle appartiennent les caractéristiques.

Définition 9.1.0.1 LDSOSTAG: *graphe étiqueté et dirigé d'adjacence spatio-temporelle de régions (LDSOSTAG, Labeled Directed Spatial Object Spatio-Temporal Adjacency Graph). Le SOTAG dirigé dont les nœuds sont étiquetés par les classes TL est appelé graphe d'adjacence temporelle dirigé et étiqueté des objets spatiaux.*

Dans ce cas, un STO, c'est-à-dire un chemin du LDSOSTAG est caractérisé par une séquence de symboles. On appelle cette séquence *séquence de caractéristiques discrètes du STO*. Un exemple de LDSOSTAG est fourni figure 9.1. Les numéros correspondent aux identifiants des nœuds, et les lettres à leurs étiquettes, c'est-à-dire aux symboles associés aux centroïdes des classes TL correspondantes.

Nous redéfinissons ici les termes classiquement utilisés dans le domaine de la recherche de motifs fréquents (précédemment définis au chapitre 6), et nous les adaptions à nos données : le LDSOSTAG.

Définition 9.1.0.2 Motif: *Un motif est un ensemble de symboles.*

Définition 9.1.0.3 Longueur d'un motif: *On appelle longueur d'un motif le nombre de symboles qui le constituent.*

L'ensemble des nœuds du graphe est noté \mathcal{U} , et les nœuds du graphe sont notés u_i , pour $i = 1 \dots |\mathcal{U}|$. Les étiquettes des nœuds du graphe correspondent à des classes TL et sont notées $\Psi^c(u_i)$. On note $s_{i,\mathcal{G}}$ les sous-chemins du LDSOTAG, et $\mathbb{S}_{\mathcal{G}}$ l'ensemble des sous-chemins. Les chemins sont définis de façon unique, ainsi, $\forall (s_{1,\mathcal{G}}, s_{2,\mathcal{G}}) \in \mathbb{S}_{\mathcal{G}}^2$, si $\forall (i, j) \in \llbracket 1, |s_{1,\mathcal{G}}| \rrbracket \times \llbracket 1, |s_{2,\mathcal{G}}| \rrbracket$, $s_{i,1,\mathcal{G}} = s_{j,2,\mathcal{G}}$ alors $s_{1,\mathcal{G}} = s_{2,\mathcal{G}}$.

On distingue deux types de sous-chemins : les sous-chemins de LDSOSTAG, et les sous-chemins de STO. Ces deux notions diffèrent car les sous-chemins de STO ne sont pas uniques dans le graphe. En effet, prenons le sous-chemin de graphe de la figure 9.1, $s_{i,\mathcal{G}} = (u_1, u_{27}, u_{53})$. Ce sous-chemin de graphe est unique dans le LDSOSTAG. En revanche, si l'on considère les STO_{13} , et STO_{14} , ces deux STO présentent comme sous-chemin (u_1, u_{27}, u_{53}) . On voit donc qu'il n'y a pas unicité des sous-chemins de STO dans le graphe.

Définition 9.1.0.4 Fréquence d'un motif dans le LDSOSTAG: *On appelle fréquence d'un motif, m , le nombre de sous-chemin de graphe dont m est la séquence des caractéristiques qui lui sont associées.*

Définition 9.1.0.5 Motif fréquent dans le LDSOSTAG: *Un motif fréquent (dans le LDSOSTAG) est un motif correspondant à au moins σ_{\min} sous-chemins du graphe.*

Par exemple, en reprenant le LDSOSTAG de la figure 9.1, et en fixant $\sigma_{\min} = 2$, $\{F, I, U\}$ est un motif fréquent de fréquence 2 et de longueur 3 : il apparaît sur les STO 2 et 5. Les chemins correspondants sont les séquences de nœuds :

$$\begin{aligned} 6, 15, 24, 37, 50, 61 & \quad (9.6) \\ 1, 12, 21, 34, 47, 58 & \end{aligned}$$

Les ensembles de symboles correspondants sont :

$$\begin{aligned} B, F, I, M, O, U & \quad (9.7) \\ A, F, I, L, P, U & \end{aligned}$$

Le graphe une fois étiqueté, on y cherche les motifs fréquents pour représenter une évolution.

On note o_i les chemins maximaux (STO) du LDSOTAG, $i \in \llbracket 1, |o| \rrbracket$, o est l'ensemble des STO du LDSOTAG, et $|o|$ est son cardinal, c'est-à-dire le nombre de STO de la STIS. Ces chemins maximaux sont des séquences de nœuds (SO) notées : $o_i = (o_i^0, \dots, o_i^{T-1})$, où T est le nombre d'images de la STIS. La séquence de classes TL correspondant à un chemin o_i est notée $\Psi^c(o_i) = (\Psi^c(o_i^0), \dots, \Psi^c(o_i^{T-1}))$. Un motif m est un sous-ensemble de $\Psi^c(o_i)$ correspondant à un sous-chemin s_i du STO o_i tel que $\Psi^c(s_i) = m$. $m = \left(\Psi \left(o_i^{\varphi^{-1}(k)} \right) \right)_{k \in \llbracket 0, l-1 \rrbracket} = (m_0, \dots, m_{l-1})$, φ étant une fonction qui renomme l'ensemble des indices du sous-ensemble en un ensemble d'indices consécutifs et croissants. Par exemple, $(m_0, m_1, m_2, m_3) = (A, F, I, U) = (\Psi(o_5^0), \Psi(o_5^1), \Psi(o_5^2), \Psi(o_5^3))$.

$$\begin{aligned} \varphi(0) & \rightarrow 0 & (9.8) \\ \varphi(1) & \rightarrow 1 \\ \varphi(2) & \rightarrow 2 \\ \varphi(5) & \rightarrow 3 \end{aligned}$$

On pourra aussi préciser la longueur d'un motif ou d'un sous-chemin en la notant en exposant. Ainsi, un motif de longueur l est noté m^l . Notons que la différence entre les sous-chemins de STO et les sous-chemins du LDSOTAG est que les sous-chemins de STO sont définis par rapport aux STO alors que les p sont définis par rapport au graphe. En conséquence, les sous-chemins de STO ne sont pas uniques : on peut avoir $i \neq j$, et $s_{k_i} = s_{k_j}$. Considérons par exemple les sous-chemins de STO $s_{i_8} = (o_8^0, o_8^4, o_8^5)$ et $s_{i_9} = (o_9^0, o_9^4, o_9^5)$ du graphe de la figure 9.1. Ces deux sous-chemins sont égaux et correspondent à l'unique sous-chemin de graphe, $s_{i,G}(u_1, u_{44}, u_{55})$. On note N_c le nombre de classes TL de la STIS.

Notons que les STO peuvent avoir une intersection non vide dans le graphe, et en conséquence, des STO différents peuvent avoir des sous-chemins égaux. Par exemple, les STO 8 et 9. Il s'agit de STO différents ayant néanmoins une intersection non vide, en conséquence le sous-chemin (o_1, o_{44}, o_{55}) est un sous-ensemble des deux STO, mais correspond au même sous-chemin dans le graphe.

Enfin, on notera \mathcal{M} l'ensemble des motifs (et \mathcal{M}^l l'ensemble des motifs de longueur l), \mathcal{F} , l'ensemble des motifs fréquents (et \mathcal{F}^l l'ensemble des motifs fréquents de longueur l).

9.1.1 Fréquence d'un motif dans le LDSOTAG

Considérons le graphe de la figure 9.1. La fréquence, ¹ du nœud A (bleu marine) seul est 1 cependant, le nombre de chemins dans le LDSOTAG orienté contenant le symbole A est 12. Une première conséquence est que la fréquence d'un nœud n'est plus le cardinal de l'ensemble des identifiants de STO, comme c'était le cas des bases de données linéaires (les STO du graphe correspondent aux objets d'une base de données linéaire). Et une deuxième conséquence est qu'un motif de longueur l peut être plus fréquent qu'un de ses motifs couvrants. La propriété d'antimonotonie qui est une des propriétés sur laquelle s'appuie un grand nombre d'algorithmes de recherche de motifs fréquents n'est donc plus vérifiée.

9.1.2 Les méthodes de recherche de motifs fréquents

Nous avons vu qu'étant données les données, nous avons choisi une analyse temporelle des données par recherche de motifs fréquents dans le graphe. Plusieurs approches existent pour ce type de problème (voire annexe ??). Nous avons néanmoins choisi d'adapter une méthode de recherche de motifs fréquents dans les bases linéaires à notre graphe, et particulièrement, l'algorithme ECLAT. Dans cette section, nous discutons des autres méthodes mentionnées en annexe ?? afin de justifier notre choix.

Nous mentionons au §A.6.1.1, une méthode de recherche de sous-graphes fréquents. Ce problème est néanmoins différent de notre problème. En effet, les structures à plusieurs branches dans le graphe d'adjacence temporelle traduisent une information de voisinage spatial, ce qui n'est pas notre centre d'intérêt ici. Les structures linéaires du graphe sont les structures qui permettent de décrire les évolutions des caractéristiques radiométriques. Notons que le fait que l'on cherche des ensembles de symboles, et non pas des sous-graphes permet d'utiliser un ordre total pour les sous-ensembles permettant d'assurer qu'un k-ensemble soit généré à partir de la jonction d'une unique paire de (k-1)-ensembles fils d'un même nœud dans l'arbre de recherche (autrement dit ayant un même préfixe).

On peut donc penser au problème de recherche de chemins fréquents dans un graphe représentant un site web pour l'analyse des séquences de clics. Dans notre cas, néanmoins,

1. Nous utilisons l'acceptation du terme fréquence comme le nombre d'occurrences.

le graphe considéré est étiqueté, et ce sont les fréquences des séquences d'étiquettes associées à des sous-chemins qui nous intéressent. C'est donc directement le graphe que nous souhaitons analyser, alors qu'au §A.6.1, le graphe était une contrainte supplémentaire à considérer pour l'analyse d'une base de données. Dans notre cas, en revanche, si l'on doit considérer une base de données, c'est le graphe lui-même. D'autre part, le graphe d'un site web contient des cycles ce qui complexifie son analyse car en conséquence, il contient des chemins infinis. Notre problème ne présente pas cette difficulté puisque notre graphe ne comprend aucun cycle. Nous voyons donc que ces problèmes qui peuvent paraître proches à première vue sont en fait très différents, et les algorithmes utilisés dans ce domaine ne s'appliquent donc pas ici.

Le problème le plus proche reste donc la recherche de motifs fréquents dans un ensemble d'objets.

Nous avons vu au ch.6 qu'une structure efficace pour le stockage et le parcours des motifs fréquents est l'arbre de préfixe.

Il y a deux façons de construire cet arbre : en largeur, ou en profondeur. L'approche en largeur permet un élagage plus important que l'approche en profondeur et est donc plus rapide. En contrepartie, cette approche requiert un grand espace de stockage car l'arbre complet doit être gardé en mémoire. Dans notre cas, l'arbre a au plus T niveaux et T vaut plusieurs dizaines. L'espace requis dans ce cas pour une recherche en largeur est trop important. Nous choisissons donc une approche en profondeur.

Nous avons présenté au ch.6 l'algorithme ECLAT [Zaki, 2000c, Zaki et al., 1997] qui est une méthode efficace de recherche en profondeur que l'on propose d'appliquer ici au cas des graphes.

9.1.3 ECLAT pour les graphes

L'algorithme ECLAT [Zaki, 2000c, Zaki et al., 1997] propose une recherche des motifs fréquents par un parcours en profondeur de l'arbre de préfixes. Afin de réduire la recherche, l'algorithme exploite la propriété d'anti-monotonie selon laquelle un sous-motif d'un motif fréquent est nécessairement fréquent. Ainsi, les motifs candidats de longueur l peuvent être obtenus en effectuant la jonction de motifs de longueur $l - 1$ différant que d'un motif. L'arbre de préfixe permet d'effectuer cette opération de façon efficace car les extensions possibles pour un nœud correspondent à ses frères à droite.

Dans le cas des graphes, nous avons vu que cette propriété n'est plus vérifiée. En revanche, un motif dont les sous-motifs sont fréquents est fréquent. L'arbre de préfixe ne contient donc généralement pas l'ensemble des motifs fréquents. En effet, l'ordre dans lequel les symboles sont ajoutés à un motif influe sur le motif final obtenu. Ainsi, en reprenant l'exemple de la figure 9.1, si le nœud A est traité en premier, l'arbre de préfixe ne contiendra aucun motif contenant A . En effet, si l'on considère le graphe dans sa totalité, la fréquence de A est 1. En revanche, si l'ordre considéré le place à la fin, des motifs fréquents le contiendront, par exemple : AMS ou ACJ .

Lors de la construction de l'arbre, les extensions d'une classe c ne se limitent donc pas aux classes plus grandes que c , mais toutes les classes qui n'appartiennent pas au chemin allant de la racine jusqu'à c sont des extensions candidates. Dans ce cas, l'espace de recherche est le treillis complet, et il n'est pas envisageable de le parcourir, surtout pour les longueurs de STO considérées.

N'ayant pas trouvé d'heuristique appropriée pour ordonner les classes et s'affranchir ainsi du problème de non monotonie, nous proposons d'effectuer un post-traitement consistant

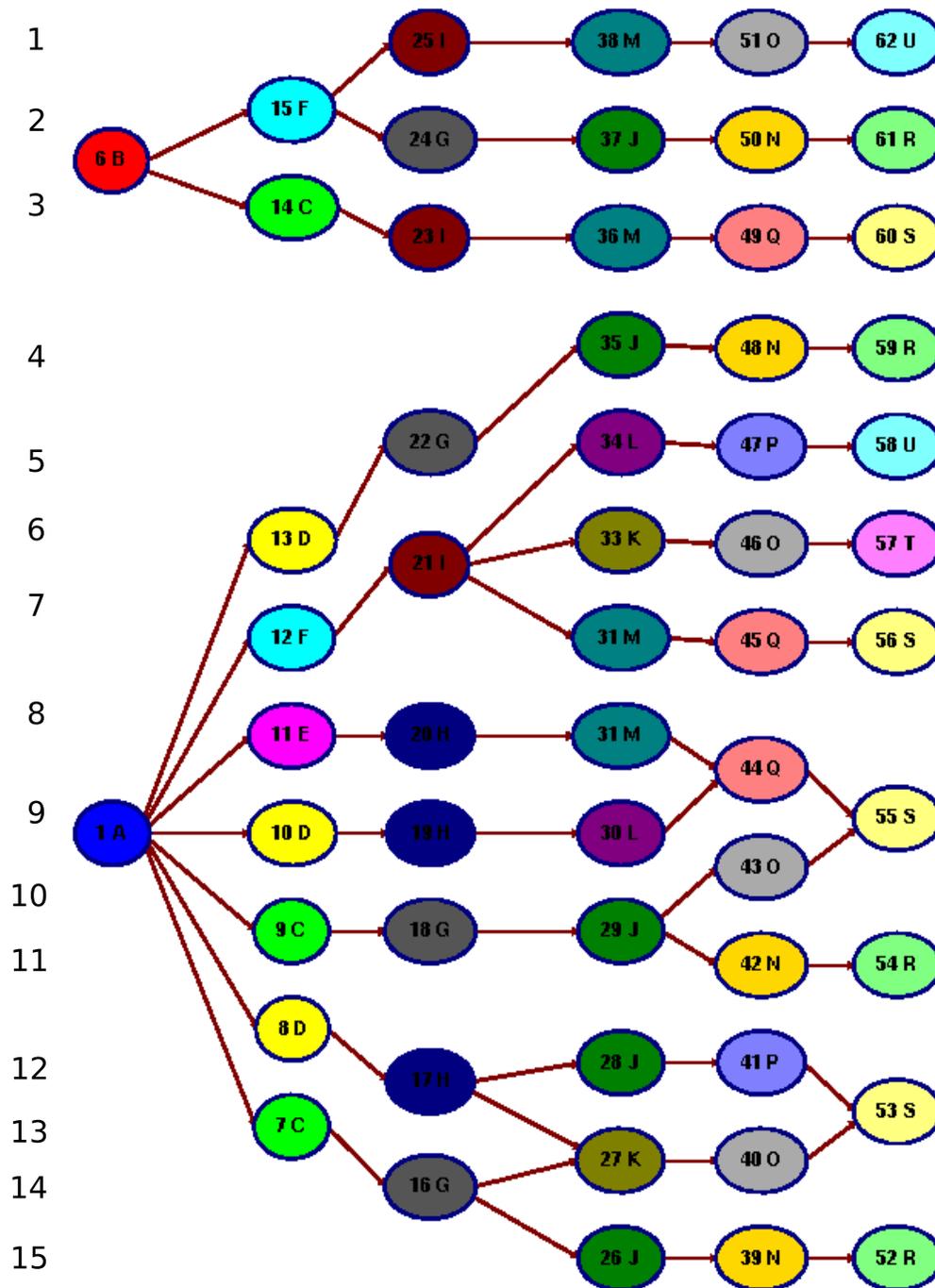


FIGURE 9.1 – Illustration d'un graphe, vu comme une base de donnée de graphes. Les numéros dans les nœuds correspondent aux identifiants de nœuds, et les lettres correspondent aux étiquettes associés (représentant un groupe de radiométries), et les numéros à gauche représentent les identifiants des chemins (c'est-à-dire des STO).

à agrandir les motifs fréquents de l'arbre de préfixe. Une façon d'agrandir un motif est d'effectuer sa fermeture. C'est-à-dire d'agrandir un motif par l'intersection des séquences de caractéristiques des STO qui le contiennent : $G(m) = \bigcap_{i \in K(m)} \Psi^c(o_i)$. Par exemple, pour le motif KO , $K(KO) = 13, 14, 4$, et la fermeture est donc $(ADHKOS) \cap (ACGKOS) \cap (AFIKOT) = AKO$. Cependant, plusieurs motifs peuvent avoir la même fermeture et une telle approche trouvera donc plusieurs fois le même motif. En effet, d'une part, si l'extension concerne des classes plus grandes, les branches correspondantes seront construites plus loin dans la construction de l'arbre, et d'autre part, tous les sous-motifs générateurs d'un motif fermé produiront le même motif par fermeture. Par exemple, pour $\sigma_{\min} = 3$, IM est un motif fermé. La fermeture de I et de M est IM . Afin de trouver tous les motifs fréquents, il nous faut donc pour chaque motif fréquent trouvé dans l'arbre, calculer sa fermeture, supprimer les classes qui sont plus grandes que la dernière classe (elles seront traitées ultérieurement), et vérifier que les motifs ainsi obtenus n'ont pas déjà été reportés comme fréquents. Il nous faut donc stocker tous les motifs fréquents, et les parcourir pour chaque ajout de motif agrandi. Cette approche est beaucoup trop coûteuse, d'une part, en temps de calcul, et d'autre part en espace de stockage.

Nous proposons donc de restreindre les motifs d'intérêt aux motifs maximaux. Pour chaque feuille de l'arbre de préfixe, on cherche sa fermeture, et on vérifie qu'elle n'a pas déjà été stockée. Il est donc nécessaire de stocker les motifs maximaux. Le nombre de motifs maximaux étant considérablement réduit par rapport au nombre de motifs fermés et a fortiori de motifs fréquents, cette approche est justifiée.

L'efficacité de l'algorithme ECLAT est principalement due à l'implémentation par bit-map permettant d'obtenir la fréquence d'un motif par des simples & logiques. En effet, la fréquence d'un motif correspond au cardinal de sa couverture. Ceci n'est malheureusement plus la cas pour le LDSOTAG. Ainsi, la fréquence du symbole A de la figure 9.1 est 1 alors que sa couverture est $\{4, 5, \dots, 15\}$. Voyons comment adapter ce principe aux graphes.

Considérons l'ensemble des sous-chemins correspondant au motif m . Un motif m correspond à un ensemble de sous-chemins du LDSOTAG, $\{s_{i,\mathcal{G}}(m)\}_{i=1}^{\sigma(m)}$ définis de façon unique. Et chacun des $s_{i,\mathcal{G}}(m)$ correspond à un ou plusieurs sous-chemins de STO. Par exemple, en considérant le motif de longueur 1, C , celui-ci a pour fréquence $\sigma(\{C\}) = 3$, les sous-chemins de graphe correspondants sont :

$$\begin{aligned} s_{0,\mathcal{G}}(m)(\{C\}) &= u_{14} \\ s_{1,\mathcal{G}}(m)(\{C\}) &= u_9 \\ s_{2,\mathcal{G}}(m)(\{C\}) &= u_7 \end{aligned} \tag{9.9}$$

Le nœud numéro 14 n'appartient qu'au STO_3 alors que les nœuds u_9 et u_7 appartiennent respectivement aux STO_{10} et STO_{11} , et STO_{13} et STO_{14} . Nous proposons de regrouper les sous-chemins de graphe, $s_{i,\mathcal{G}}(m)$ correspondant à un seul sous-chemin de STO dans un ensemble appelé *ensemble de chemins indépendants de LDSOTAG* ou ensembles indépendant de LDSOTAG, et ceux auxquels correspondent plusieurs sous-chemins de STO dans un ensemble appelé *ensemble de chemins dépendants de LDSOTAG* ou ensemble dépendant de LDSOTAG :

$$\begin{aligned} \{s_{i,\mathcal{G}}(m)\}_{i=1}^{\sigma(m)} &= I^{\mathcal{G}}(m) \cup D^{\mathcal{G}}(m) \text{ avec } I^{\mathcal{G}}(m) \cap D^{\mathcal{G}}(m) = \emptyset \\ I^{\mathcal{G}}(m) &= \{s_{i,\mathcal{G}}(m) \mid \Psi^{\mathcal{G},c}(p(m)) = m \text{ et } \exists ! o_j \mid p_{k_j} = s_{i,\mathcal{G}}(m)\} \\ D^{\mathcal{G}}(m) &= \left\{ s_{i,\mathcal{G}}(m) \mid \Psi^{\mathcal{G},c}(p(m)) = m \text{ et } \exists \{o_j \mid p_{k_j} = s_{i,\mathcal{G}}(m)\}_{j=1}^n \text{ et } n > 1 \right\} \end{aligned} \tag{9.10}$$

On appelle *ensemble indépendant de STO de m* , et on note $I(m)$ l'ensemble des identifiants de STO correspondants à $I^{\mathcal{G}}(m)$, et on appelle *ensemble de groupes dépendants de STO de m* , et on note $D(m)$ l'ensemble des groupes d'identifiants de STO correspondants à $D^{\mathcal{G}}(m)$.

On peut stocker l'ensemble indépendant de STO de m dans un vecteur de bits dont les colonnes représentent les STO, et l'ensemble des groupes dépendants dans une matrice de bits dont les lignes correspondent aux différents groupes, et dont les colonnes correspondent aux identifiants de STO. Par exemple, en reprenant le graphe de la figure 9.1, pour le motif G , on a alors :

$$\begin{array}{l}
 I(G) = \begin{array}{cccccccccccccccc}
 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 \\
 \begin{array}{|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|}
 \hline
 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 \hline
 \end{array} \\
 \\
 D(G) = \begin{array}{cccccccccccccccc}
 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 \\
 \begin{array}{|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|}
 \hline
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\
 \hline
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\
 \hline
 \end{array}
 \end{array}
 \end{array}$$

Les $s_{i,\mathcal{G}}(m) \in \{s_{i,\mathcal{G}}(m)\}_{i=1}^{\sigma(m)}$ sont disjoints (par l'unicité des sous-chemins de LDSOTAG). Or chacun d'eux correspond soit à un sous-chemin de STO indépendant, soit à un groupe de sous-chemins de STO dépendants. Donc les lignes de la matrice de bits représentant l'ensemble des groupes dépendants ont une intersection vide. De plus, par définition de la décomposition en ensemble indépendants et en ensemble de groupes dépendant (l'eq.9.10), $I^{\mathcal{G}}(m) \cap D^{\mathcal{G}}(m) = \emptyset$, donc le vecteur de bit correspondant à l'ensemble indépendant d'un motif a une intersection vide avec l'ensemble des vecteurs lignes de la matrice des groupes dépendants. La figure 9.2 représente les matrices I et D du niveau 0 de l'arbre de préfixe du graphe de la figure 9.1.

La jonction de deux motifs m_1 et m_2 nécessite d'effectuer les quatre intersections entre ensembles dépendants et indépendants des deux motifs. On définit l'intersection entre une matrice de bits et un vecteur de bits comme étant l'intersection de toutes les lignes de la matrices avec le vecteur, et on définit l'intersection entre deux matrices de bits comme étant l'intersection entre chaque ligne de la première matrice vers chacune des lignes de la deuxième matrice.

L'intersection d'un ensemble indépendant avec un autre ensemble indépendant ou un ensemble dépendant est une sous-partie de l'ensemble indépendant de la jonction des motifs, et l'intersection entre les ensembles de groupes dépendants des motifs se décompose en une sous-partie de l'ensemble indépendant de la jonction et l'ensemble de groupes dépendant de la jonction :

$$I(m_1) \cap I(m_2) = I_1(m_1 \cup m_2) \tag{9.11}$$

$$I(m_1) \cap D(m_2) = I_2(m_1 \cup m_2) \tag{9.12}$$

$$D(m_1) \cap I(m_2) = I_3(m_1 \cup m_2) \tag{9.13}$$

$$D(m_1) \cap D(m_2) = I_4(m_1 \cup m_2) \cup D(m_1 \cup m_2) \tag{9.14}$$

$$D(m_1 \cup m_2) = \bigcup_{i=1}^4 I_i(m_1 \cup m_2) \tag{9.15}$$

Les ensembles indépendants de l'intersection des ensembles dépendants des deux motifs s'obtiennent en sélectionnant les lignes de la matrice de dépendance qui ne contiennent qu'un seul 1.

	I															D															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
C	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	
D	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0
G	0	1	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	
H	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	
I	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	
J	0	1	0	1	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0
M	1	0	1	0	0	0	1	1	0	0	0	0	0	0	0																
N	0	1	0	1	0	0	0	0	0	0	0	1	0	0	0	1															
O	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	
Q	0	0	1	0	0	0	1	1	0	0	0	0	0	0	0																
R	0	1	0	1	0	0	0	0	0	0	1	0	0	0	1																
S	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	
																0	0	0	1	0	0	0	0	0	0	0	1	1	1	0	

FIGURE 9.2 – Matrices I et D du niveau 0 de l'arbre de préfixe du graphe de la figure 9.1.

Etant donné que les lignes des matrices d'ensembles dépendants sont disjointes, l'intersection d'une ligne de $D(m_1)$ est non vide avec au plus une ligne de $D(m_2)$. Les matrices n'ont donc pas besoin d'être parcourues systématiquement complètement pour calculer l'intersection.

Nous avons vu que le support d'un motif pouvait être calculé simplement à partir de ses matrices de bits représentant les ensembles de groupes dépendants et les vecteurs de bits représentant les ensembles indépendants. De plus nous avons décrit comment obtenir ces ensembles pour un motif à partir de sous-motifs. L'algorithme ECLAT [Zaki, 2000c, Zaki et al., 1997] peut donc être appliqué aux graphes.

Cet algorithme consiste à calculer récursivement pour chaque nœud la base de données conditionnelle consistant dans le cas des bases de données linéaires à une matrice dont les lignes correspondent aux 1-extensions fréquentes, et les colonnes aux identifiants de transactions supportant le motif étendu. Dans notre cas, la base de données conditionnelle consiste donc en la matrice dont les vecteurs lignes sont les ensembles indépendants de chaque extension, et en une liste de matrices correspondant aux ensembles de groupes dépendants de chaque extension. Le calcul de la base de données conditionnelle d'une extension s'obtient donc par les quatre intersections des équations 9.11 à 9.15 entre les différents vecteurs lignes de la matrice d'indépendance et matrices dépendantes.

9.1.4 Ordre des motifs

Afin de réduire le nombre de candidats, on peut, comme pour les bases de données linéaires, ordonner les classes TL par fréquence croissante. Un tel réordonnement des

classes est possible grâce au codage choisi pour le graphe selon lequel les identifiants de STO sont connus pour chaque classe. Une recherche des motifs directement dans le graphe pour le calcul des fréquences des motifs comme c'est le cas dans une approche Apriori ne permettrait pas un tel réordonnement.

9.2 Interprétation théorique des évolutions fréquentes

Nous avons présenté une méthode permettant de retrouver efficacement les motifs fréquents maximaux dans le LDSOTAG. Ces motifs capturent l'information structurée présente dans les évolutions de caractéristiques. L'interprétation sémantique de ces motifs n'est cependant pas toujours claire, et nous proposons dans ce paragraphe une interprétation théorique dans le cas d'évolutions typiques de NDVI.

Considérons différentes évolutions de NDVI de différentes cultures. Idéalement, on souhaite retrouver les signaux regroupés par classes. Pour différentes raisons, la durée temporelle des motifs fréquents peut être réduite. Afin d'aider l'interprétation des résultats et le réglage des paramètres longueur et forme, nous proposons donc une étude de ces différentes perturbations.

9.2.1 Superposition

Nous avons supposé que la scène observée ne contenait que des objets de l'arrière-plan. Cependant, cette hypothèse n'est pas toujours vérifiée. En particulier, des nuages superposent partiellement ou complètement certains objets de la scène. En fonction du nombre de nuages à un instant, la classe nuage seule sera supposée fréquente et apparaîtra donc dans l'arbre de préfixe ou pas. Notons que le nombre de nuages dans une image dépend d'une part de la nature de la scène, mais aussi de la segmentation.

Dans le cas où la classe TL *nuage* n'est pas fréquente, on peut se poser la question de savoir si cette classe TL sera ajoutée lors de la phase d'agrandissement des motifs comme c'est le cas des zones de terre nues. L'étape de fermeture consiste en une intersection de tous les STO contenant le motif fréquent à agrandir, or autant ils contiennent tous des phases de terre nue, autant, ils n'ont pas été tous superposés par un nuage au même instant. Les classes TL de nuages n'étant pas fréquentes seules ne seront donc pas ajoutées aux motifs.

Dans le cas de scènes particulièrement nuageuses, de nuages composites ou de sur-segmentation des nuages, la classe TL correspondante peut être trouvée fréquente. Plusieurs configurations sont alors possibles :

- Les nuages sont dispersés sur la scène mais recouvrent peu souvent les mêmes objets. Dans ce cas, les motifs contenant le nuage ne contiendront pas plus d'une classe TL.
- Les nuages (composites ou simples) recouvrent des zones composites de même classe TL (c'est le cas par exemple des villes). Dans ce cas, des motifs fréquents et longs contenant des nuages pourront être trouvés.
- Les nuages sont composites et plusieurs petits nuages recouvrent une même zone. Dans ce cas encore, des motifs fréquents et longs contenant des nuages pourront être trouvés.
- Les nuages recouvrent des zones composites de classe TL différentes (c'est le cas par exemple d'un nuage superposant plusieurs parcelles). Dans ce cas, les motifs contenant le nuage ne contiendront pas plus d'une classe TL.

Notons que même dans le cas où des motifs contenant un nuage sont extraits, il existe des sur-motifs du motif sans nuage qui sont plus fréquents.

9.2.2 Effet des variations d'évolutions intra-classes, et quantification vectorielle

En fonction de la terre, et de l'organisation des agriculteurs, les STO d'une même classe peuvent évoluer légèrement différemment, induisant des décalages temporels, et des variations dans les gradients temporels. Ces variations ont pour effet d'augmenter la disparité intra-classe lors des phases de gradient temporel important des caractéristiques, réduisant donc l'écart entre la disparité intra-classe et la disparité inter-classe. Deux classes ayant des phases de transition proches pourront donc se trouver regroupées par la classification TL. En conséquence la plage temporelle de différenciation de différentes cultures est réduite. De plus, lors des phases de transition, en fonction de la répartition des décalages, et du gradient des évolutions, la quantification pourra regrouper des parties de classes, séparant ainsi les SO d'une même classe d'évolution en plusieurs groupes. De plus, la frontière entre classes TL évolue en fonction du temps, et les sous-groupes ainsi formés ne sont pas nécessairement durables.

9.2.3 Effet du croisement d'évolutions

Les évolutions de caractéristiques peuvent se croiser, et en fonction de la configuration spatiale des STO qui les supportent, ces intersections pourront résulter en une augmentation de la fréquence. Notons que pour une intersection donnée, plus il y a de classes TL se croisant, plus il y a de chance que les SO correspondants soient fusionnés aussi dans l'espace image. Ainsi, en début et en fin de cycle, lorsque toutes les parcelles correspondent à de la terre nue, il n'en résulte pas nécessairement une augmentation de la fréquence de la classe terre nue (voire même une baisse) car les STO correspondant sont aussi fusionnés dans l'espace image. Les zones de croisements ne concernent cependant généralement qu'un nombre réduit d'instant, et plus un motif sera long, moins il aura de chances de correspondre à une zone de croisement.

9.2.4 Effet d'erreurs de segmentation ou d'appariement

Il existe finalement une troisième source d'erreurs liée à la nature séquentielle de l'algorithme. Certaines erreurs de segmentation auront pour effet de sur-segmenter des SO. En conséquence, la fréquence de la classe TL correspondante sera légèrement surestimée. Cette erreur ne gêne pas la recherche de motifs. Réciproquement, des erreurs de segmentation peuvent avoir pour effet de sous-segmenter certains SO. La fréquence de la classe TL correspondante sera légèrement sous-estimée. Cette erreur n'est pas non plus gênante si les mises en correspondance entre les instants précédents et suivants sont correctement effectués, car dans ce cas, soit la fréquence de la classe TL reste au dessus de σ_{\min} et rien n'est changé, soit la fréquence passe en dessous, et dans ce cas, la classe TL sera ajoutée lors de la phase d'agrandissement des motifs. Enfin des erreurs de segmentation peuvent aussi avoir pour effet une mauvaise localisation de la frontière dont une conséquence peut être une erreur d'appariement.

Les erreurs d'appariement sont les plus graves. En effet, d'une part, elles complexifient le graphe ce qui peut rallonger considérablement le temps de calcul, mais d'autre part, l'ajout d'un seul *faux* arc engendre au moins un *faux* chemin, et avec lui (ou eux), tous ses (leurs) sous-chemins. On pourra néanmoins réduire leur effet en augmentant le seuil de fréquence.

9.2.5 Regroupement par prototypes

Ces différentes erreurs tendent à augmenter le nombre de valeurs divergentes. Afin de prendre en compte ce point, nous avons considéré des motifs de support temporel non compact. Cependant ces valeurs divergentes n'ont aucune raison d'apparaître aux mêmes instants pour les différents STO d'un même type. En conséquence, d'une part, les motifs sont moins fréquents, et le seuil de fréquence ne devra donc pas être trop élevé, mais d'autre part, il en résulte une multiplication du nombre de motifs fréquents. Par ailleurs, l'approche gloutonne que nous avons adoptée nous empêche de prendre en compte une décision plus souple que l'égalité lors du calcul des fréquences. Nous proposons donc d'effectuer ces regroupements lors d'un post-traitement.

La quantification est effectuée directement dans l'espace multi-canal. En conséquence, on ne peut comparer les symboles de façon cohérente. Nous proposons donc une autre approche pour le regroupement des motifs.

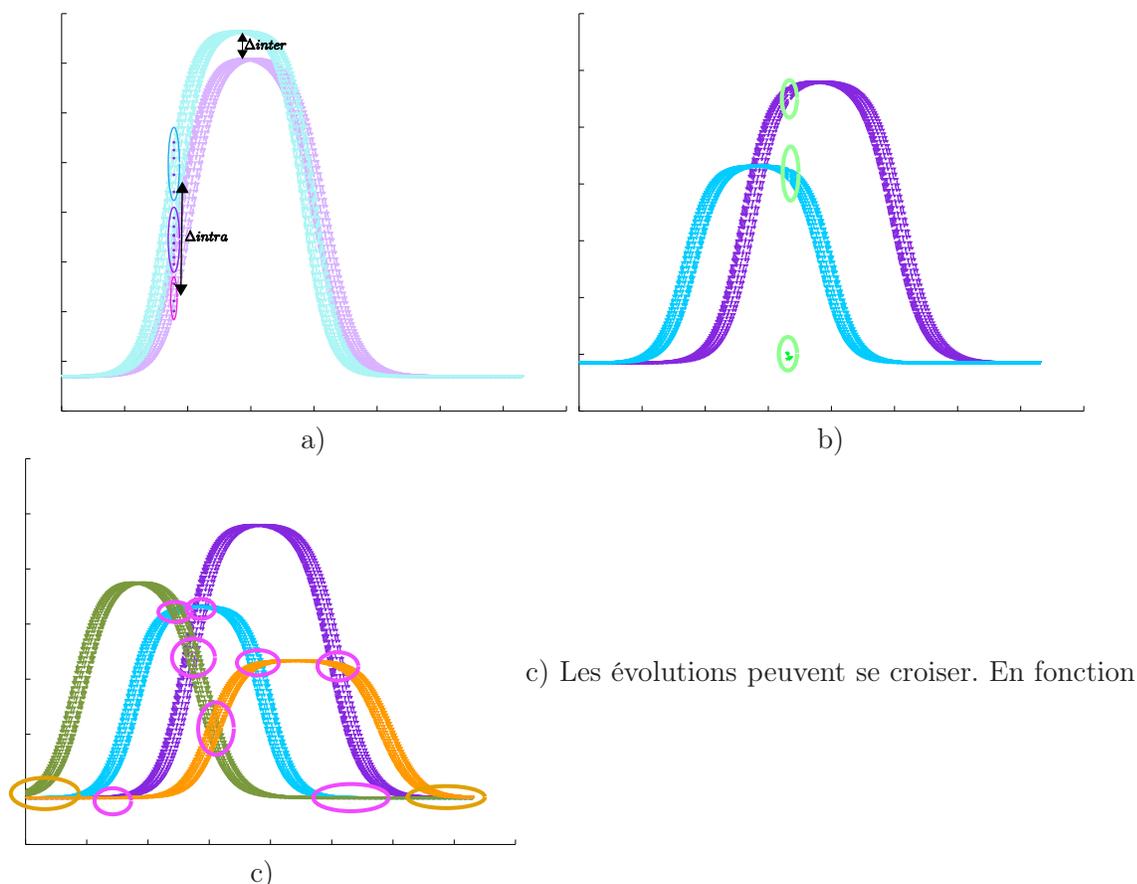
Remarquons que beaucoup des motifs fréquents de l'arbre de préfixes ne diffèrent que d'une classe TL, mais à des instants différents. Afin de regrouper ces motifs, nous proposons de regrouper les motifs en-dessous d'un certain seuil de distance de Hamming par rapport à un *prototype*.

Pour ce faire, on considère une structure supplémentaire contenant l'ensemble des prototypes que l'on stocke lors de la génération des motifs fréquents. Au moment d'insérer un motif fréquent maximal, on parcourt l'ensemble des prototypes jusqu'au premier avec lequel la distance de Hamming est inférieure à un seuil qu'on s'est donné. Si un tel prototype est trouvé, alors, on incrémente simplement son compteur. Si aucun prototype suffisamment proche n'est trouvé, alors on crée un nouveau prototype égal au motif en question.

Ce post-traitement peut être effectué avant ou après l'agrandissement des motifs par fermeture. Les agrandissements sont sensés capturer les parties communes des classes d'évolution et doivent donc être appliqués aux groupes les plus peuplés possibles. En effet, dans le cas contraire, des extensions spatialement locales risqueraient d'avoir lieu, typiquement l'insertion de nuages. Afin d'éviter ces extensions néfastes, nous effectuons le regroupement par prototypes avant l'agrandissement.

Cependant, l'agrandissement nécessite les identifiants des STO supportant le motif. La structure contenant les prototypes doit donc aussi stocker les identifiants. Comme précédemment, nous proposons de les stocker dans des octets codant des vecteurs de bits.

Détaillons l'agrandissement des motifs après regroupement par prototypes. Cette opération a lieu en tout dernier lieu, après avoir complètement parcouru l'arbre de préfixe. Pour chaque prototype $s_{i,g}$, l'intersection de l'ensemble des STO stockés dans le vecteur de bits b_i est d'abord calculée. Pour ce faire, un simple $\&$ des lignes correspondantes dans la matrice de bits représentant l'ensemble des STO comme une base de données linéaire est effectuée. On capture ainsi, les éventuelles extensions du motif. En revanche, la partie correspondant au motif a toutes les chances d'être vide car les différents STO regroupés dans un même prototypes peuvent présenter des classes TL différentes en différentes positions. Afin de récupérer le motif complet, un *ou* est donc finalement calculé entre la fermeture du motif et le motif lui-même.



c) Les évolutions peuvent se croiser. En fonction

de la topologie dans l'espace image, des croisements d'évolution auront pour effet d'augmenter la fréquence des motifs. Ainsi, les croisements entourés d'une ellipse rose auront moins de chance de correspondre aussi à une fusion dans l'espace image car ils couvrent moins de régions. Ceux qui sont entourés d'une ellipse orange, en revanche auront beaucoup plus de chances de correspondre à une fusion aussi dans l'espace image, et en conséquence à ne pas être fréquents.

FIGURE 9.3 – a) Le décalage a pour effet que la disparité intra-classe au niveau de forts gradients temporels est du même ordre que la disparité inter-classe à maturation. Ainsi, $\Delta_{intra} \sim \Delta_{inter}$. En fonction de la densité de la classe, ces décalages peuvent donner lieu à des divisions intra-classe lors de la quantification vectorielle. C'est par exemple le cas pour le groupement violet qui regroupe des valeurs de la classe MT bleue et de la classe MT mauve. b) Un nuage a pour effet de regrouper des objets éventuellement différents dans l'espace de caractéristiques. En fonction de la densité du nuage, il correspondra dans l'espace image à une ou plusieurs régions. Dans le cas où il correspond à une région, son effet par rapport à la fréquence des motifs superposés est nul puisque dans ce cas, la fréquence est nulle. Dans le cas où il est multiple dans l'espace image, le motif TL sera donc perçu comme fréquent, ce qui correspondra, dans le cas où des SO de natures différentes sont occultés, à un regroupement de ces objets à tort. Cependant, en raison du seuil de longueur minimale considéré, les sur-motifs de ce motif seront non fréquents.

9.3 Réglage de la longueur et de la fréquence des motifs pertinents

Lorsqu'on effectue une classification, on cherche une partition de l'espace vectoriel considéré. La description d'un espace par une partition est intéressante de plusieurs points de vues. D'une part, elle permet une description complète des données. D'après le principe du MDL, taux-distortion ou encore MAP, un nombre optimal de classes est obtenu pour le meilleur compromis entre la simplicité du modèle, et la fidélité des données décrites par le modèle. Un tel compromis n'est possible que si le modèle fournit une description complète des données. Et d'autre part, les partitions sont caractérisées par la non interpénétrabilité des composantes. Cette propriété réduit l'espace de recherche lors de la construction de telles partitions. Par conséquent, des algorithmes mettant en compétition les différentes composantes peuvent être appliqués : nuées dynamique, EM, etc.

En considérant une description des comportements radiométriques des STO de la STIS par des évolutions à support non convexe et à longueur variable, on perd la propriété de complétude du modèle et par conséquent, le paramétrage automatique, et la non-interpénétrabilité des composantes.

Nous avons vu qu'un motif pertinent doit être long, et fréquent. La description n'étant pas complète, les seuils correspondants ne peuvent être optimisés par un critère informatif. Afin de permettre une description complète des données par le modèle afin de pouvoir les comparer, on considère que les valeurs n'appartenant pas à un motif fréquent d'évolution sont des valeurs divergentes. On suppose alors que ces valeurs sont les réalisations d'un processus aléatoire suivant une distribution uniforme. Pour les autres caractéristiques, celles appartenant aux motifs fréquents d'évolution, elles peuvent être décrites, soit à partir d'un modèle d'évolution temporel en supposant une calibration acceptable puisqu'elles sont maintenant dénuées de ces valeurs intempestives, soit par une séquence de caractéristiques TL. Il existe cependant un grand nombre de possibilités pour la description d'une séquence. En effet, chaque séquence est décomposable selon k partitions, et chaque composante de ces partitions correspond à une combinaison des symboles de la séquence. Chaque configuration correspond à une certaine longueur de description de la séquence ainsi que des autres séquences contenant les motifs impliqués. Il en résulte une optimisation dans un espace de recherche prohibitif. C'est pourquoi nous avons opté pour un réglage de ces paramètres manuel utilisant des heuristiques issues des considérations d'interprétation des motifs que nous venons d'effectuer.

9.3.1 Conclusion

Les motifs fréquents extraient l'information structurée de la séquence et sont donc pertinents au sens informationnel. Leur interprétation est néanmoins difficile, et nous avons présenté, dans cette section, un exemple d'interprétation théorique des motifs fréquents fournis par la méthode.

Ce type de considération laisse penser que des approches d'analyse par raisonnements basés sur des règles constituent une approche prometteuse pour l'interprétation de la scène grâce à des motifs fréquents. Ces raisonnements peuvent s'appuyer principalement sur des ontologies faisant intervenir les évolutions des caractéristiques des objets (comme c'est le cas de l'interprétation que nous avons présentée), mais elle peut aussi s'appuyer sur des propriétés géométriques. Ainsi, nous avons vu différentes configurations correspondant à une scène nuageuse, en fonction des changements topologiques observés. Cependant, l'in-

interprétation que nous avons effectuée reste théorique, et une validation des raisonnements sur des données réelles reste à faire.

Une autre approche serait d'attacher une sémantique aux motifs obtenus par apprentissage. Cependant, les motifs sont de taille variable, et de plus, une classe sémantique correspond souvent à une disjonction de plusieurs motifs. Par exemple, dans le cas où certains motifs d'une classe ont été altérés par une des perturbations sus-mentionnées. Or les algorithmes classiques d'apprentissage sont linéaires et ne permettent donc pas d'apprendre des disjonctions. Les algorithmes considérant une hiérarchie de l'information sémantique permettent cependant de prendre en compte des disjonctions. Ce type d'algorithme d'apprentissage pourrait donc être appliqué pour attacher une sémantique à ces motifs.

Cinquième partie

Expérimentations et résultats

Chapitre 10

Validation

10.1 Segmentation spatiale sous contrainte temporelle

L'algorithme proposé est un algorithme itératif, et dépend donc d'une initialisation. Nous allons dans un premier temps discuter l'initialisation et présenter la méthode adoptée, puis, nous présenterons les résultats.

10.1.1 Prétraitements

Afin de simplifier le problème, nous avons considéré une simple correspondance entre les OS. Cependant, en réalité les OS sont généralement liés à plusieurs OS, et une sélection de l'OS correspondant est donc effectuée. Par ailleurs, l'optimisation que nous avons choisie est de type descente de gradient. Ces deux points ont pour conséquence qu'une mauvaise décision concernant le choix du meilleur correspondant d'un OS aura des répercussions graves sur les partitions finales obtenues. Il est donc important, qu'au moment de l'initialisation, les correspondances 1-1 sélectionnées correspondent effectivement à des paires de régions incluses dans des régions mises en correspondance dans la séquence de partitions optimale. Afin d'assurer la pertinence des correspondances 1-1 choisies à l'état initial, nous initialisons la séquence de partitions par une séquence de partitions triviales,¹ pour lesquelles les correspondances sont évidentes.

La matrice de covariance associée à une région de taille 1 est nulle, ce qui entraîne des coûts de fusion infinis d'après 5.12. Afin de permettre des fusions avec des régions de taille unitaire, nous adoptons la même approximation que dans [Kanungo et al., 1994] qui considère une moyenne pondérée de la matrice de covariance locale et de la matrice de covariance globale pour les petites régions (en-dessous d'un seuil n_0).

La conséquence de cette pondération est que les régions de taille unitaire ont une matrice de covariance proportionnelle à la matrice de covariance de l'image, et tendent donc à se fusionner avec des régions très hétérogènes. Ainsi, une région a tendance à agglomérer toutes les régions de un pixels voisines et à grossir ainsi jusqu'à atteindre la taille de l'image. [Kanungo et al., 1994] propose un pré-traitement qui consiste à effectuer d'abord toutes les fusions entre régions unitaires avant de considérer des fusions avec des régions de deux pixels. Pour ce faire, les meilleures fusions sont effectuées, mais les fusions impliquant la région résultante ne sont pas réinsérées dans la pile de fusions candidates, et de plus, toutes les fusions candidates impliquant une région voisine de cette région sont supprimées.

1. chaque région correspond à un pixel.

Dans notre cas, une telle approche n'est pas applicable puisque rien n'assure que les mêmes pixels soient fusionnés dans les mêmes images, et donc la correspondance entre régions serait perdue. Nous proposons donc un pré-traitement au cours duquel des cylindres de régions unitaires sont fusionnés.

Imposer des cylindres de la longueur de la séquence est très contraignant, et peu de fusions sont effectuées dans ce cas. Aussi, nous avons introduit un autre paramètre W_p permettant de régler la longueur des cylindres qui fusionnent lors du pré-traitement. Dans nos expériences, nous avons choisi $W_p = W$.

La distance entre un cylindre, $c_1^{t_0}$, et son cylindre voisin, $c_2^{t_0}$, est le maximum sur tous les instants du cylindre des distances radiométriques : $D(c_1^{t_0}, c_2^{t_0}) = \max_{t \in \llbracket t_0, t_0 + W_p \rrbracket} \|c_1^{t_0}(t) - c_2^{t_0}(t)\|_2$. Pour chaque cylindre, la distance avec les 4 cylindres voisins (en connexité 4) est calculée, et si la distance minimale sur ces 4 cylindres voisins est inférieure à un seuil s_p , alors les cylindres sont fusionnés :

$$c_1^{t_0} \text{ est fusionné (avec son voisin } c_2^{t_0}) \Leftrightarrow \min_{c_2^{t_0} \text{ Vois } c_1^{t_0}} \left(\max_{t \in \llbracket t_0, t_0 + W_p \rrbracket} \|c_1^{t_0}(t) - c_2^{t_0}(t)\|_2 \right) \leq s_p \quad (10.1)$$

10.1.2 Résultats

Nous proposons dans un premier temps d'étudier le comportement de l'algorithme sur des cas simulés. Les séquences considérées présentent des conservations, des fusions, une superposition, et petites déformations géométriques. Nous avons choisi une tolérance aux déformations égales aux déformations maximales présentes dans les images, c'est-à-dire un pixel pour les figures 10.3 et 10.4, et de deux pixels pour 10.1. De plus, nous avons choisi des paramètres radiométriques proches pour deux régions voisines stables temporellement de façon à mettre en évidence l'avantage de la segmentation jointe par rapport à une séquence de segmentations indépendantes. Ainsi, nous voyons dans la figure 10.1, que là où des segmentations indépendantes ne parviennent pas à séparer deux régions de radiométries très similaires mais temporellement stables, notre algorithme parvient à capturer cette stabilité malgré de petites déformations des régions. De même, dans la figure 10.3, on observe que notre algorithme permet de capturer la stabilité des régions malgré des décalages de frontières, et un changement topologique. Enfin, dans 10.4, nous validons la robustesse de l'algorithme aux occlusions.

Un des objectifs de la méthode de segmentation que nous proposons est de réduire le nombre de paramètres à régler. Cependant, nous avons vu que le codage que nous proposons est redondant. En conséquence, la mise à échelle n'est pas forcément optimale telle quelle. Nous proposons donc d'introduire un paramètre, comme c'était le cas dans [Delyon et al., 2006].

L'équation 7.13 devient alors :

$$\mathcal{L} = \sum_{r \in \text{Im}_1, \dots, \text{Im}_T} (1 - \lambda)(\mathcal{L}_{\text{geom}}^{\text{ST}}(r) + \mathcal{L}_{\text{param}}(r)) + \lambda \mathcal{L}_{\text{radiom}}(r) \quad (10.2)$$

Notons que l'introduction de ce paramètre n'est pas gênante outre mesure car seul un paramètre de finesse des segmentations est à régler, les finesses s'ajustant automatiquement entre les différentes segmentations ce qui est nettement moins laborieux que le réglage

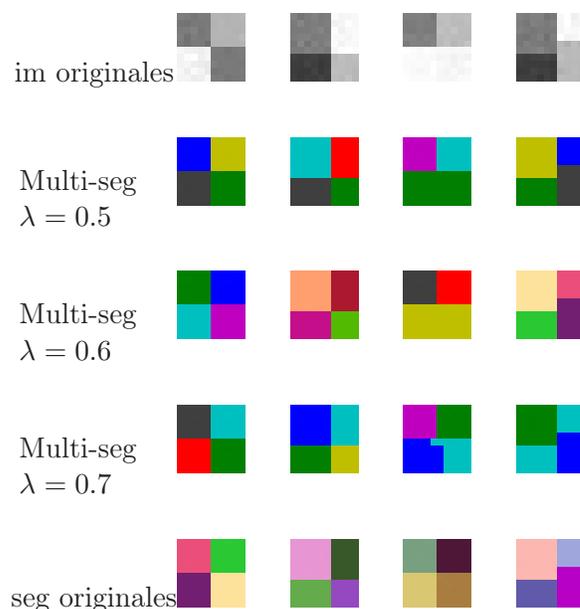


FIGURE 10.1 – Fusion avec des petits décalages. La première ligne représente les images simulées. Les deux régions similaires de la troisième image ont une différence entre leurs valeurs moyenne égale à 3. Les trois lignes suivantes présentent les résultats de segmentation multi-temporelle jointe pour trois valeurs de λ . Enfin, la dernière ligne présente les partitions utilisées pour générer les images de la première ligne.

d'un paramètre de finesse pour chaque segmentation comme le requièrent les algorithmes classiques de segmentation.

Des résultats pour différentes valeurs de λ sont présentées dans la figure 10.1. Plus λ augmente, plus le terme d'attache aux données prend de l'importance ce qui a pour conséquence d'augmenter le nombre de régions. Le paramètre $\lambda = 0.7$ semble bien adapté au cas joint alors que, comme c'était noté dans [Delyon et al., 2006], le paramètre $\lambda = 0.5$ est mieux adapté au cas indépendant. Lorsqu'aucune précision n'est donnée concernant le paramètre λ , sa valeur est de 0.5 dans le cas indépendant, et de 0.7 dans le cas joint.

La figure 10.5a) présente les temps de calculs des multi-segmentations obtenus pour des séquences de longueurs temporelles différentes. Le temps de calcul est linéaire avec le nombre d'images de la séquence. En revanche, le temps de calcul croît très rapidement avec le nombre de régions : ainsi, pour une séquence de quatre images de taille 50×50 , le temps de calcul de la pile de fusions est dix fois celui nécessaire au calcul d'une séquence de 10 images de taille 20×20 . En 10.5b), sont présentés les temps de calcul de la pile de fusion pour différentes tailles d'images. Notons que ces temps pourraient très certainement être réduits en considérant un ordre partiel de la pile de fusion par exemple en utilisant une structure de tas ou encore en stockant un certain nombre de sous-graphes pour éviter de les recalculer.

Etant donné que le temps de calcul de l'algorithme croît avec le nombre de régions initiales, nous avons découpé la sous-séquence à traiter en sous-séquences de longueur 4 et de taille 25×25 de façon à traiter ces imagerie en parallèle. Un coût maximal négatif a été choisi de façon à ce que les segmentations restent des sur-segmentations pour éviter l'effet de bloc que peut avoir pour conséquence le découpage arbitraire de la séquence.

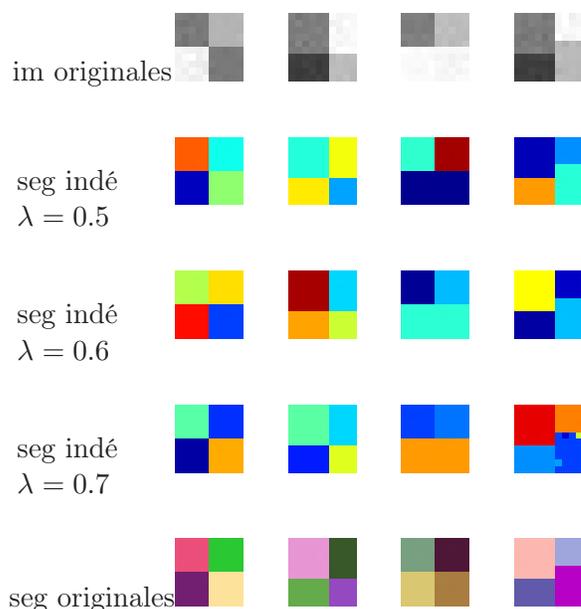


FIGURE 10.2 – Fusion avec des petits décalages. La première ligne représente les images simulées. Les régions proches de la troisième image ont une différence entre leurs valeurs moyenne égale à 3. Les trois lignes suivantes présentent les résultats de segmentation MDL indépendantes pour trois valeurs de λ . Enfin, la dernière ligne présente les partitions utilisées pour générer les images de la première ligne. Quelle que soit la valeur de λ , on ne parvient pas à distinguer les deux régions proches de la troisième image.

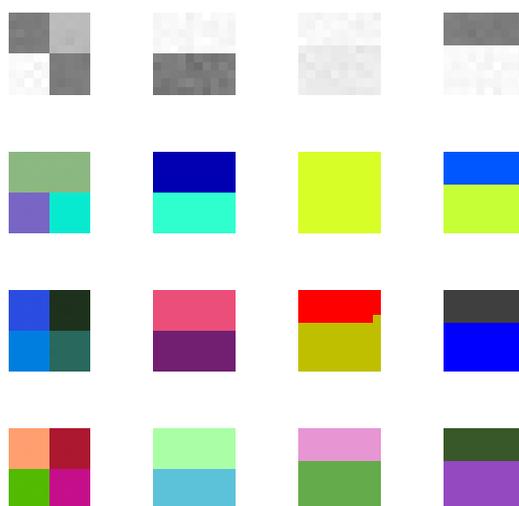


FIGURE 10.3 – Fusion avec des petits décalages. La première ligne représente les images simulées. Les régions proches de la troisième image ont une différence entre leurs valeurs moyenne égale à 6. la deuxième ligne présente les résultats de segmentation d'une segmentation MDL indépendante, la troisième ligne, le résultat de segmentation dans le cas de la segmentation multi-temporelle jointe, et la dernière ligne, les partitions utilisées pour générer les images de la première ligne.

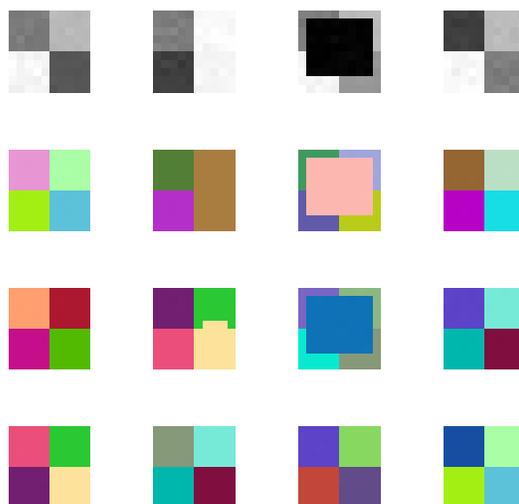


FIGURE 10.4 – Superposition avec des petits décalages. La première ligne représente les images simulées. Les régions proches de la troisième image ont une différence entre leurs valeurs moyenne égale à 4. La deuxième ligne présente les résultats de segmentation d’une segmentation MDL indépendante, la troisième ligne, le résultat de segmentation dans le cas de la segmentation multi-temporelle jointe, et la dernière ligne, les partitions utilisées pour générer les images de la première ligne.

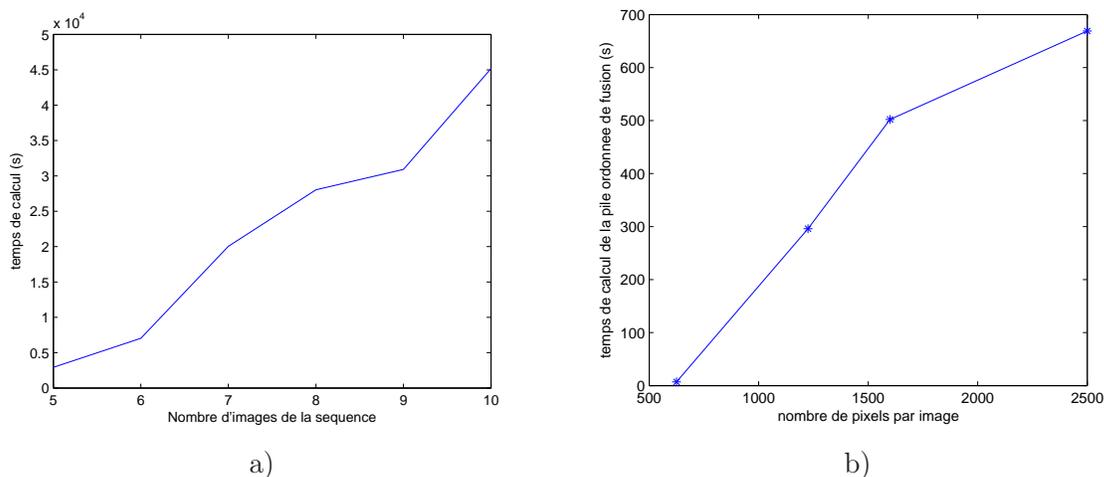


FIGURE 10.5 – Temps de calcul : a) Temps de calcul de la multi-segmentation pour des séquences de longueur variable et de taille 20×20 . b) Temps de calcul de la pile ordonnée de fusions pour des séquences d’images de taille variable et de longueur 4.

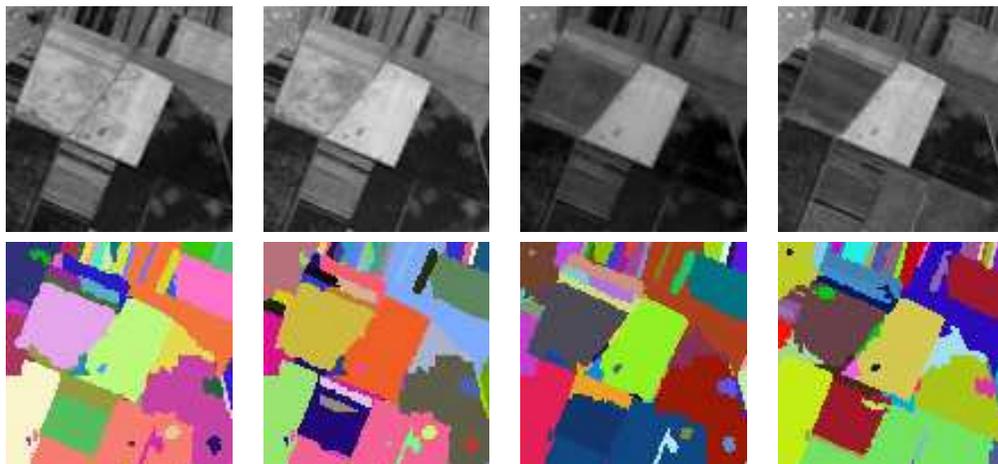


FIGURE 10.6 – Multi-segmentation jointe par partitionnement. Les segmentations obtenues sont, comme souhaité, très similaires.

Les partitions obtenues, après réassemblage, sont ensuite prises comme sur-segmentation initiale de façon à effectuer les fusions entre les régions d’imassettes différentes. Les résultats sont présentés figure 10.6.

Les figures 10.7 et 10.9 présentent les résultats pour deux séquences de quatre images de taille 25×25 . Ici encore, nous voyons qu’une valeur de $\lambda = 0.7$ paraît assez bien adaptée.

Les figures 10.8 et 10.10 sont les résultats pour les mêmes valeurs de λ dans le cas de segmentations indépendantes.

10.2 Construction du SOTAG

Les objets une fois extraits sont mis en correspondance dans un graphe d’adjacence temporelle. En raison d’erreurs de segmentation, et de petites déformations des frontières, des décisions de mise en correspondance doivent être prises. Nous avons présenté au ch.8 un algorithme de mise en correspondance itératif. Cet algorithme nécessite trois paramètres que l’utilisateur doit régler : le seuil initial, le pas de croissance de seuil et le seuil maximal. La figure 10.11 présente les résultats sur un exemple jouet présentant les différents types de changements que l’on souhaite traiter. Nous avons choisi une valeur minimale de seuil 1, un pas de 2 et un seuil maximal de 22. D’autres graphes, correspondants à des séquences simulées présentées au §10.3.1 peuvent être trouvés dans les figures 10.12, 10.17, et 10.18.

10.3 Recherche de motifs fréquents dans le SOTAG

Une fois le graphe construit, nous proposons de chercher des sous-motifs correspondants à des sous-chemins du graphe dont le comportement radiométrique est fréquent. L’algorithme est présenté ch.9.

Etant donné que nous disposons de partitions comparables issues de l’algorithme de segmentations jointes seulement pour la séquence de longueur 4 présentée figure 10.10, et que des segmentations indépendantes fournissent des graphes trop complexes, nous proposons de valider l’algorithme sur une séquence simulée dont les pixels des régions suivent une loi gaussienne pour laquelle des segmentations indépendantes sont acceptables. De

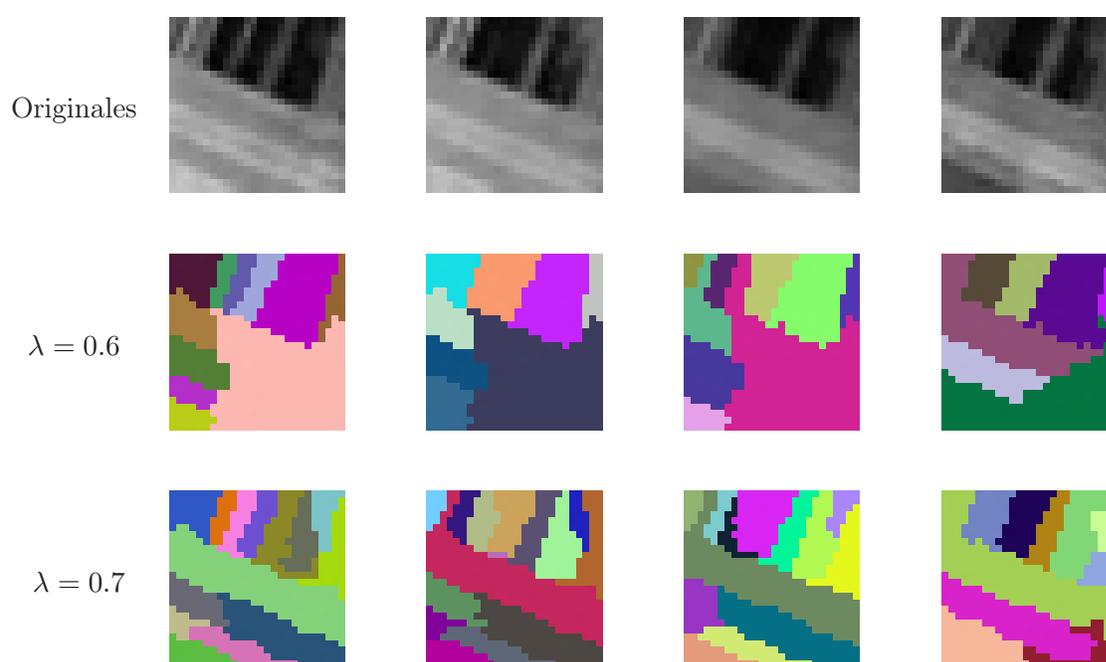


FIGURE 10.7 – **Multi segmentation jointe** : La première ligne présente les images originales, et les deux lignes suivantes, les résultats de segmentation jointe pour deux valeurs de λ . La dernière ligne est un résultat satisfaisant : malgré les faibles différences entre les radiométries des régions, l’algorithme parvient à les distinguer en exploitant leur stabilité temporelle. Les changements entre images sont généralement simples. La plupart sont des conservations, à part la région verte de la première image qui se fusionne avec une partie de sa voisine bleue pour donner la région rouge, et des fusions entre la troisième et la quatrième image. Les résultats correspondants dans le cas indépendants sont présentés figure 10.8.

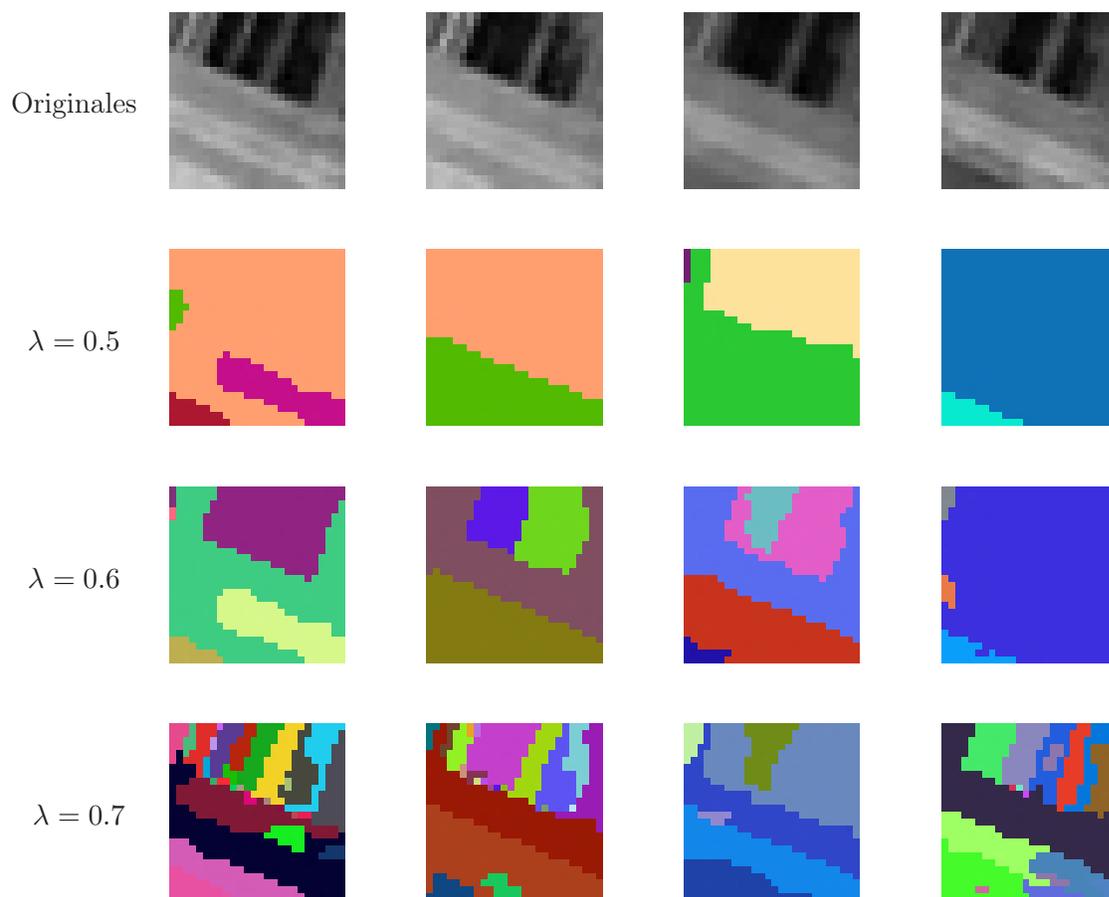


FIGURE 10.8 – **Segmentation MDL indépendante** : La première ligne présente les images originales, et les deux lignes suivantes, les résultats de segmentation jointe pour trois valeurs de λ . Ces résultats sont à comparer avec ceux de la figure 10.7. Nous avons présenté de plus les résultats pour $\lambda = 0.5$ car ce sont ceux qui sont généralement choisis.

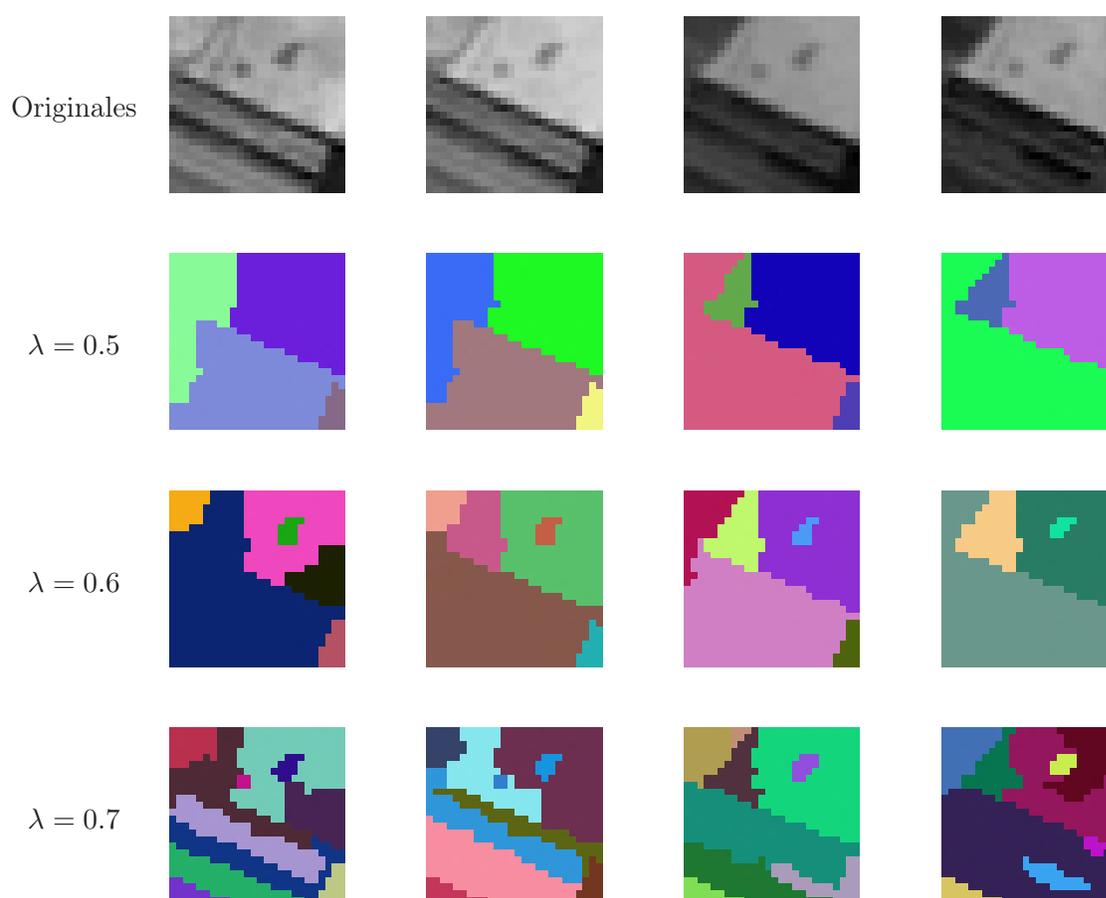


FIGURE 10.9 – **Multi segmentation jointe** : La première ligne présente les images originales, et les trois lignes suivantes, les résultats de segmentation jointe pour différentes valeurs de λ . Ici encore, l’algorithme fournit des segmentations satisfaisantes pour $\lambda = 0.7$. Les résultats sont à comparer avec ceux de la figure 10.10.

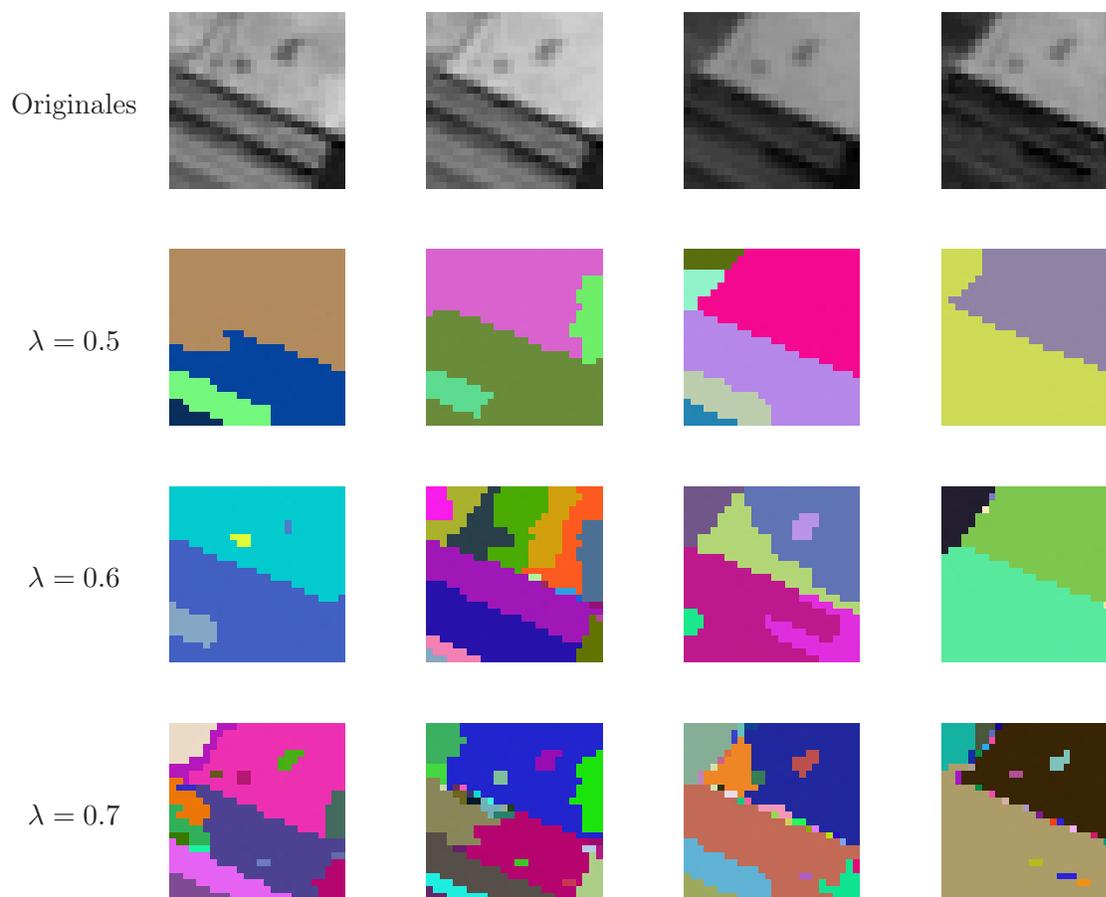


FIGURE 10.10 – **Segmentation MDL indépendante** : La première ligne présente les images originales, et les deux lignes suivantes, les résultats de segmentation jointe pour deux valeurs de λ .

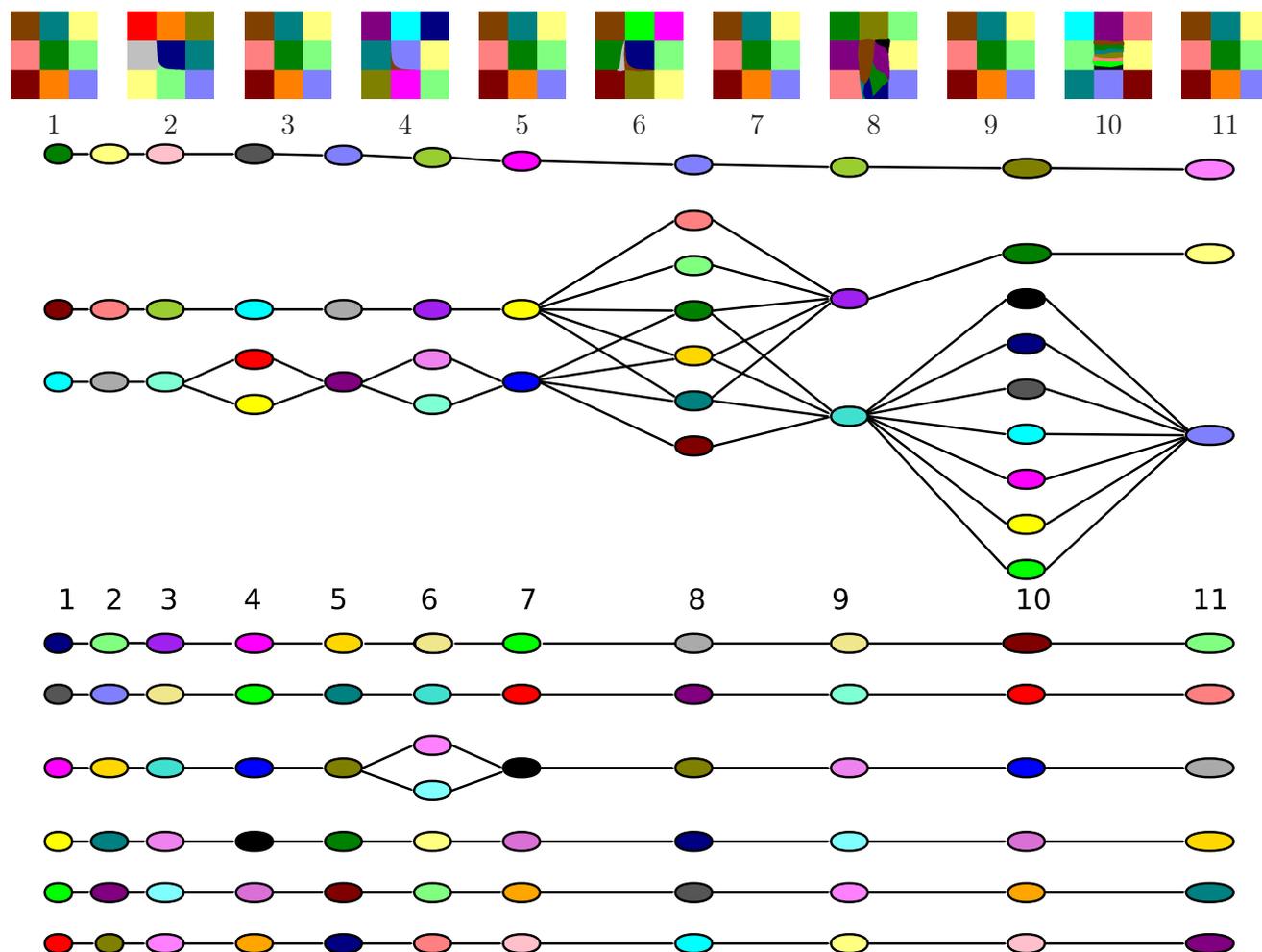


FIGURE 10.11 – La première ligne présente la séquence de partitions à partir de laquelle le graphe est construit. Cette séquence présente les différents changements types. Des changements sont présents dans les images paires par rapport aux images impaires, ces dernières étant identiques. L'image 2 présente des conservations et une erreur de localisation de frontière, l'image 4 présente une division suivie d'une fusion, l'image 6 présente deux divisions (suivies des fusions correspondantes), avec en plus de légères déformations des contours. L'image 8 présente une combinaison, et l'image 10 une division multiple avec déformations des contours. Les temps sont reportés au milieu du graphe car les nœuds ne sont pas alignés avec les images. Notons que les distances de Hausdorff entre la région bleue de l'image 2 et les régions rose et orange de l'image suivante et précédente sont plus importantes que les distances entre la région bleue (respectivement marron) de l'image 10 et la région verte de l'image 11. Pourtant, aucun arc n'existe dans le graphe pour le premier cas alors que des arcs existent dans le deuxième conformément à ce que l'on souhaite grâce à l'approche itérative permettant de faire croître progressivement les seuils de distances.

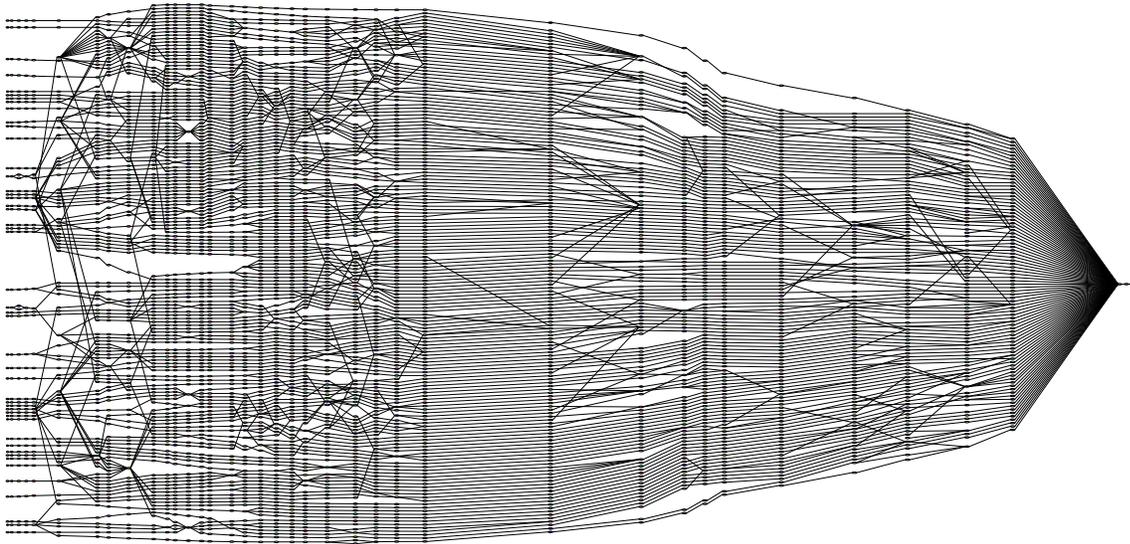


FIGURE 10.12 – Graphe associé à la séquence simulée *Fischer*. Pour la construction de ce graphe, nous avons choisi un seuil minimal de 1, un pas de 2, et un seuil maximal de 8.

plus, l'utilisation d'une telle séquence permet d'avoir une *vérité-terrain*. La génération de cette séquence sera présentée au §10.3.1. Puis, nous présenterons comment sélectionner les chemins pertinents dans le graphe, et enfin, nous présenterons les résultats obtenus.

10.3.1 Génération des séquences simulées

Dans cette section, nous présentons comment nous avons généré la séquence sur laquelle nous avons validé les algorithmes de recherche de motifs fréquents dans les graphes. Cette séquence a été obtenue en générant différents profils de culture en utilisant différents paramétrages pour la formule de Fischer [Fischer, 1994] :

$$NDVI(t) = vb + \frac{k}{1 + \exp(-c(t - p))} - \frac{k + vb - ve}{1 + \exp(-d(t - q))} \quad (10.3)$$

t est la variable correspondant au temps, $t = 0$ correspond au premier janvier, vb et ve sont des constantes correspondant respectivement à la valeur du NDVI au début et à la fin de la période de végétation, p , et q correspondent aux instants des deux points d'inflexion, et k est liée à la valeur maximale du NDVI. Le paramétrage pour les cultures d'orge, maïs, betterave, et blé est celui de [Liang, 2005]. Pour le blé, le maïs et la betterave à sucre, nous avons considéré les *cultures d'hiver*, ainsi que les *cultures d'été*. Les cultures d'hiver sont généralement plantées à l'automne, et commencent donc à pousser pendant l'hiver. À partir de la première gelée, elles cessent de croître, et demeurent dans un état de dormance jusqu'au printemps où elles reprennent leur croissance. La principale différence entre la culture d'hiver et la culture d'été est donc la présence de végétation pendant la période hivernale. Nous avons ainsi généré sept modèles d'évolution du NDVI que nous avons représenté figure 10.13 a). Pour chaque classe, on considère ensuite plusieurs échantillons. Ces différents échantillons modélisent différents champs d'une même culture et peuvent donc présenter un décalage temporel. Nous avons donc considéré un décalage temporel d'une amplitude de quinze jours en tirant aléatoirement une valeur de retard ou d'avance d'amplitude comprise entre zéro et sept jours et demi. Les évolutions des cent échantillons

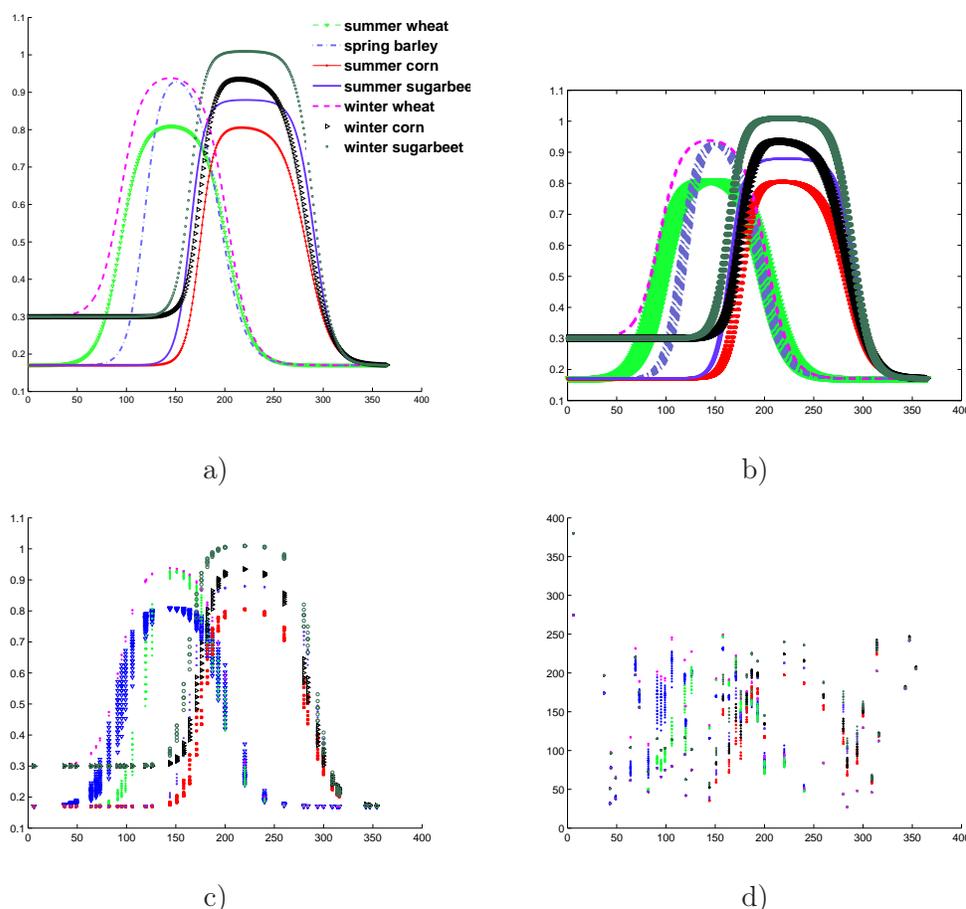


FIGURE 10.13 – a) Sept modèles d'évolution de cultures. b) Evolutions de cent échantillons de ces sept cultures avec certains décalages temporels. c) Evolutions temporellement échantillonnées. d) Evolutions perturbées par des erreurs de calibration radiométrique.

sont présentées dans la figure 10.13 b). Nous avons ensuite échantillonné temporellement ces évolutions avec un échantillonnage de *type ADAM*, c'est-à-dire que nous avons utilisé les mêmes dates d'acquisition auxquelles nous avons rajouté des dates entre le 26 juillet et le 15 octobre correspondant respectivement aux dates de fin et de début de l'acquisition. Les courbes échantillonnées sont présentées figure 10.13 c). Enfin, nous avons considéré une erreur de calibration pour chaque acquisition : deux paramètres α_t et β_t sont tirés aléatoirement et uniformément pour chaque instant t . Les points résultants sont représentés figure 10.13 d).

Nous avons ensuite généré la carte des différents types de cultures en tirant dans un damier constitué de cent cases une des cent évolutions précédemment générées pour chacune des cases. La carte correspondante est représentée figure 10.14. On génère ensuite à partir du damier des identifiants d'échantillons une séquence d'images.

Etant donné que l'algorithme de recherche de motifs fréquents est long pour une séquence de 40 images, nous avons sélectionné 15 images dans la séquence *Fischer* initiale, sur lesquelles la recherche de motif est très rapide (moins d'une seconde). La sous-séquence considérée est présentée dans la figure 10.15. De plus, afin d'étudier l'effet d'une occultation, nous avons généré une seconde séquence obtenue à partir de la première en remplaçant

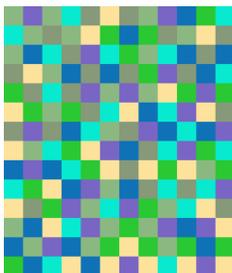
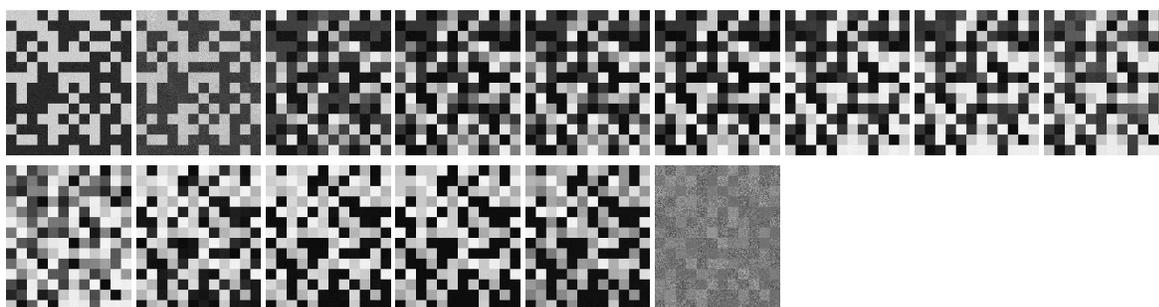


FIGURE 10.14 – Carte des évolutions

FIGURE 10.15 – Sous-séquence de 15 images de la séquence *Fischer*.

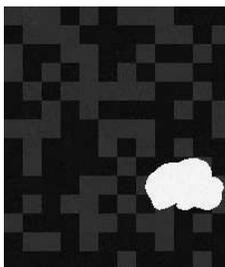
la deuxième image par celle de la figure 10.16.

10.3.2 Résultats pour les séquences simulées

Les images de la séquence sont dans un premier temps segmentées, et les valeurs de chaque image sont regroupées à l'aide de l'algorithme de regroupement de régions présenté §9.1.

10.3.2.1 Sélection des chemins pertinents

Nous avons, lors de la construction du SOTAG, sélectionné des correspondances entre images successives. Or dans cette section, nous nous intéressons non plus à des correspondances entre paires d'images, mais à des chemins traversant le graphe sur toute sa longueur. On considère un chemin comme pertinent si l'intersection de l'ensemble des régions associées aux nœuds qui le composent est non vide. Etant donné que sur une année,

FIGURE 10.16 – Image remplaçant l'image 2 dans la séquence de la figure 10.15 pour constituer la séquence dite *avec nuage*.

il y a une période pendant laquelle beaucoup de régions se divisent, et une autre, pendant laquelle beaucoup se fusionnent, il y aura beaucoup de chemins non pertinents parmi tous les chemins du SOTAG.

Par ailleurs, en superposant l'ensemble des partitions, on peut calculer l'ensemble des chemins de la longueur de la séquence ayant une intersection non vide (par exemple en multipliant les images de label). Les chemins ainsi obtenus ne sont pas non plus tous pertinents car certains peuvent correspondre à des petites déformations des régions ou à des erreurs de segmentation.

L'ensemble des chemins pertinents correspond à l'intersection de ces deux ensembles de chemins. Etant donné que l'ensemble des chemins du SOTAG est très grand, nous proposons dans un premier temps de calculer les chemins obtenus par intersection successive des partitions, puis, de sélectionner, parmi ces chemins, ceux qui sont aussi présents dans le SOTAG. Pour l'exemple jouet de la figure 10.11, le nombre de séquences fournies par le calcul de l'intersection totale est 87 et 19 d'entre eux sont supprimés car ils n'existent pas dans le SOTAG.

Les graphes de la séquence *Fischer*, de la sous-séquence sans nuage et de la sous-séquence avec nuage sont représentées respectivement figure 10.12, 10.17, et 10.18.

10.3.2.2 Résultats de l'algorithme Eclat pour les graphes

La figure 10.19 présente les longueurs et le nombre des motifs obtenus pour les deux séquences simulées. On peut constater que le nombre de motifs est restreint, ce qui traduit une bonne compression par rapport aux structures complexes des figures 10.17 et 10.18.

Les motifs peuvent être représentés en masquant la séquence par les régions correspondant aux chemins supportant les motifs. De telles séquences sont présentées figure 10.20. Les motifs présentés sont deux motifs de la deuxième séquence illustrant le bon comportement de l'algorithme dans le cas d'occultation.

Nous avons présenté dans la figure 10.21 les motifs fréquents de longueur 14 et 15 pour les deux séquences. Nous avons sélectionné une seule image par séquence par soucis de simplicité. L'image sélectionnée est celle qui présente le moins de régions. Ces images sont appelées *images-résumé*. On observe que les motifs correspondent effectivement à des régions d'une même classe.

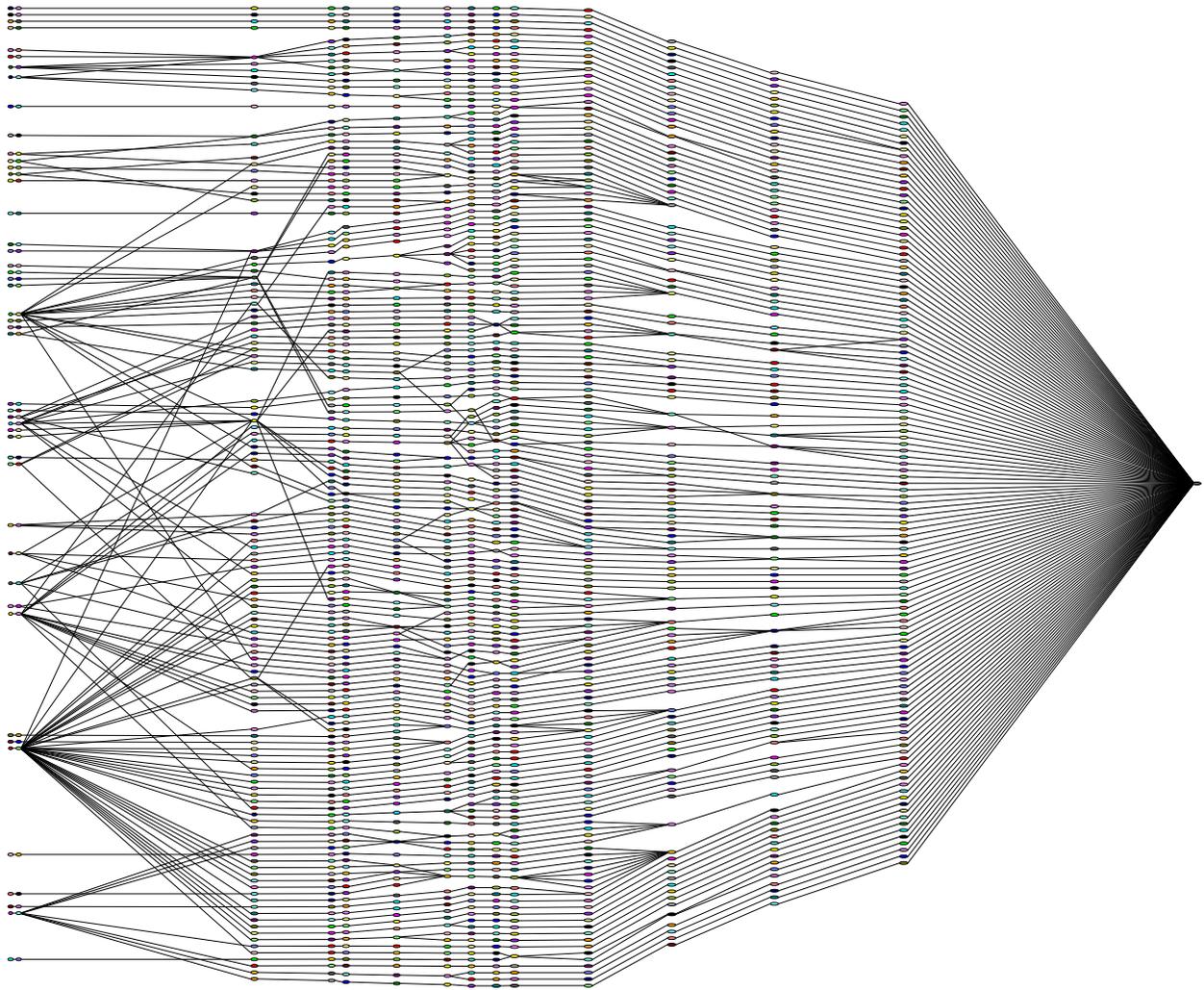


FIGURE 10.17 – Graphe correspondant à la sous-séquence sans nuage (voir figure 10.15).

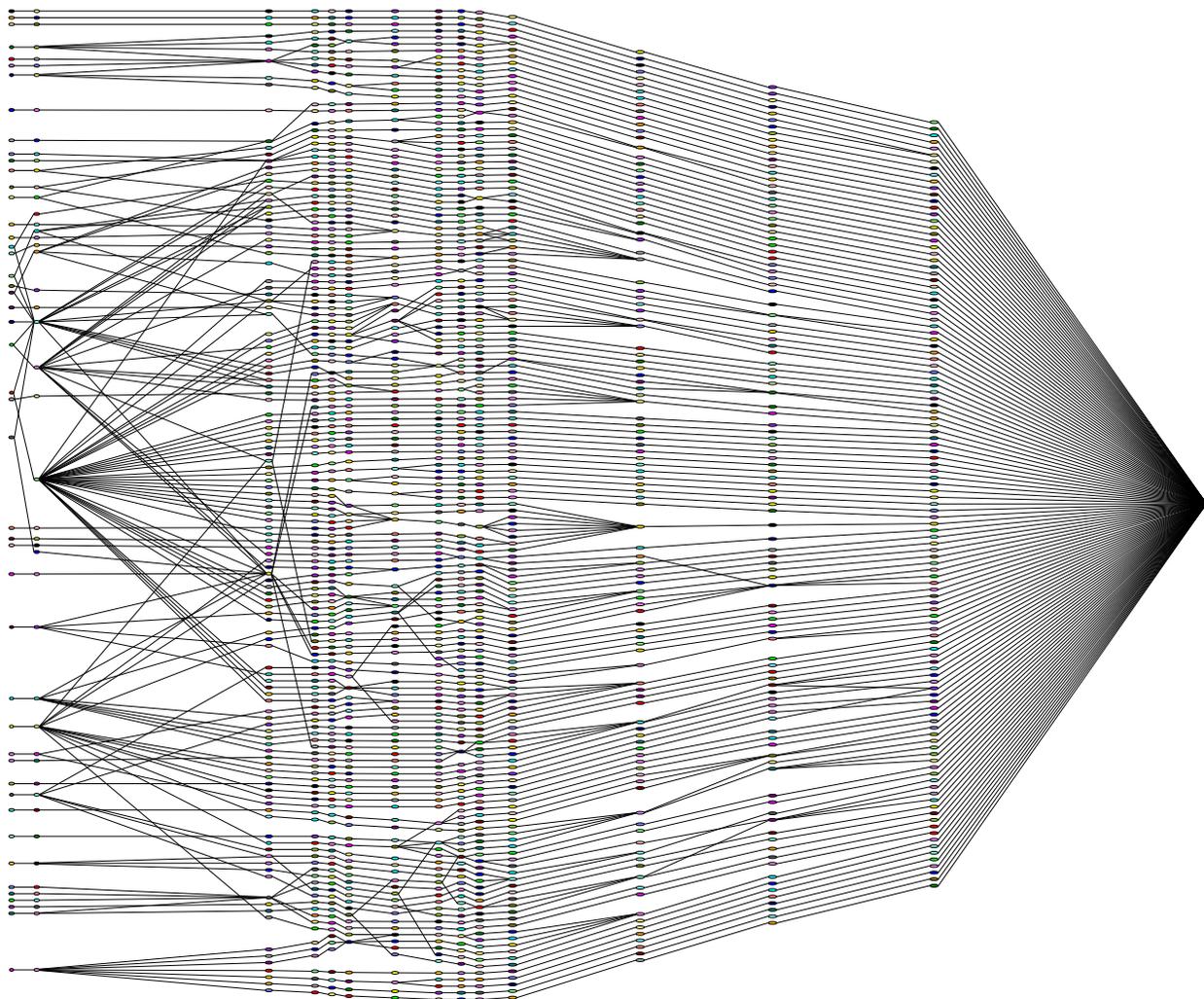


FIGURE 10.18 – Graphe correspondant à la sous-séquence avec nuage (voir figures 10.15 et 10.16).

longueur de motif	1	2	3	4	5	7	8	9	10	11	12	13	14	15
nombre de motifs dans le cas sans nuages	1	1	1	2	0	1	1	3	2	1	1	4	8	5
nombre de motifs dans le cas avec nuages	1	4	3	2	1	2	2	6	3	4	6	12	10	4

FIGURE 10.19 – Longueur des motifs obtenus.

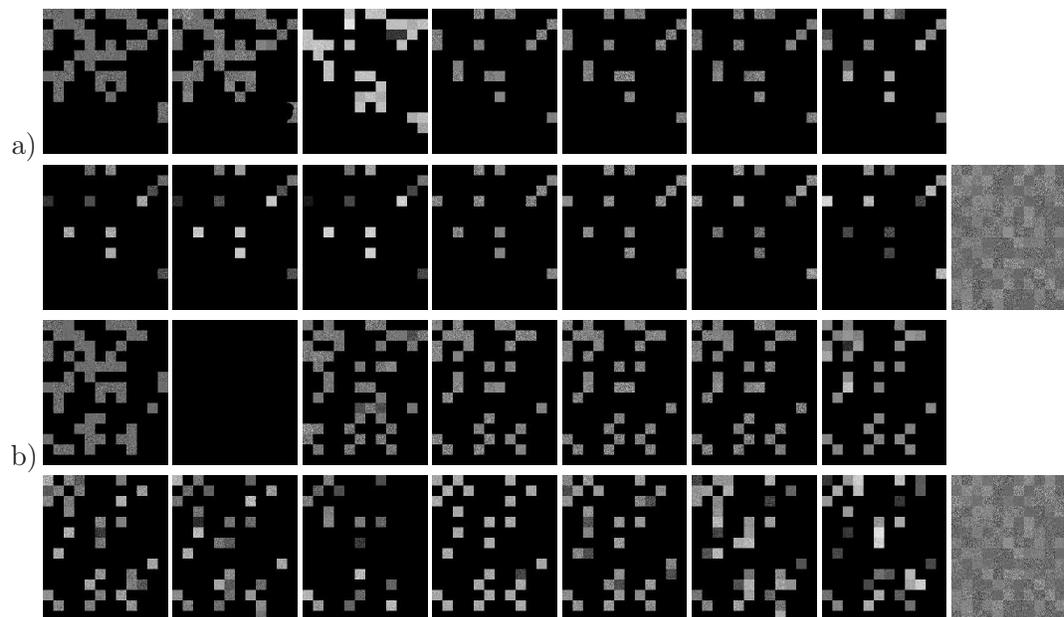


FIGURE 10.20 – Deux séquences correspondant à un motif fréquent de longueur 15 et 14 dans le cas de la séquence simulée *Fischer* avec nuage. En a), on voit en bas à droite de la deuxième image une région que le nuage superpose. L'autre chemin, passant par le nuage n'est pas sélectionné par l'algorithme comme pertinent. En b), le motif correspond à une région complètement occultée par le nuage. L'algorithme permet effectivement de retrouver le motif de support non compact dont la valeur occultée par le nuage est exclue.

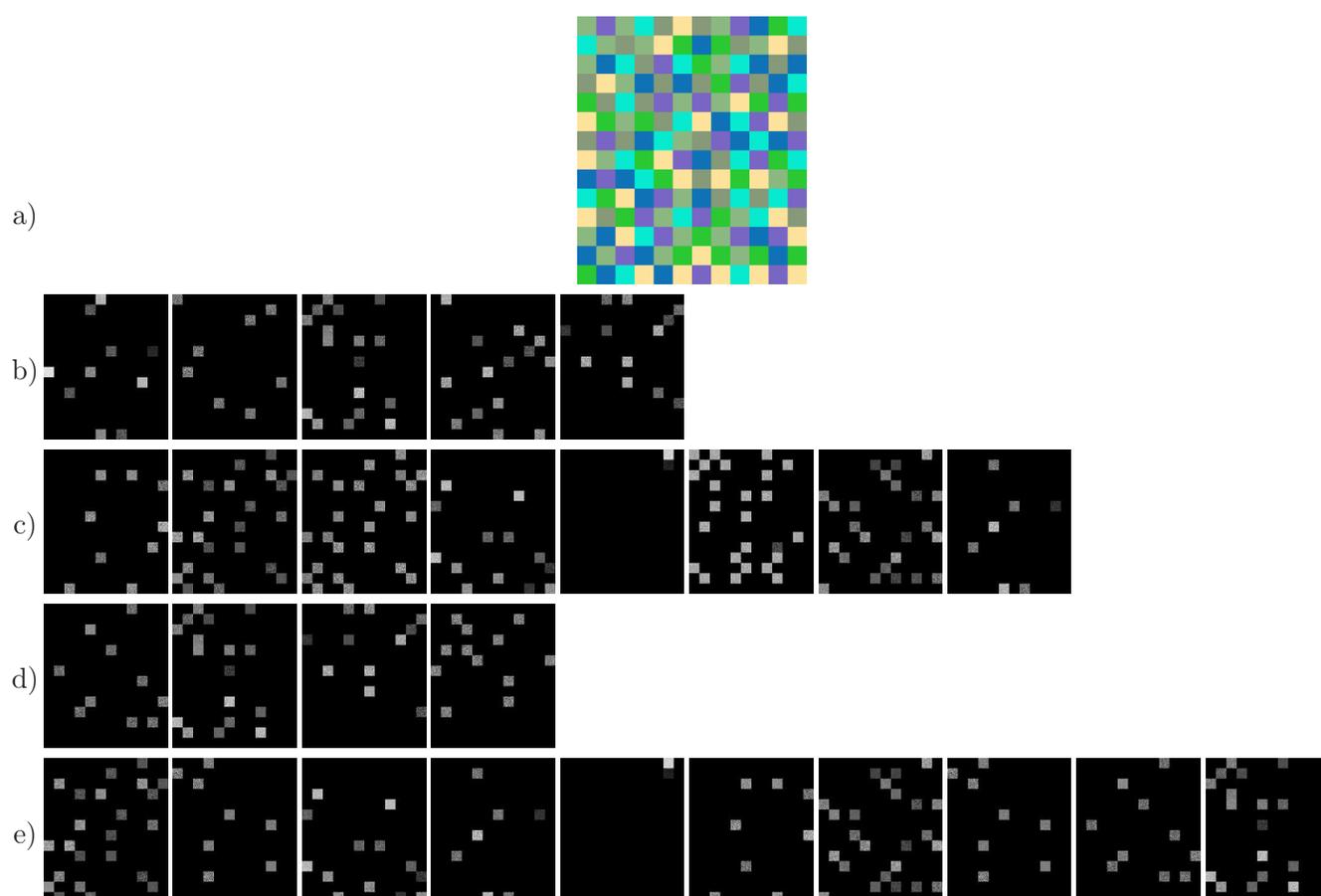


FIGURE 10.21 – a) Image des classes à partir de laquelle la séquence a été générée. b) *Images-résumé* des motifs de longueur 15 dans le cas de la séquence sans nuages. c) *Images-résumé* des motifs de longueur 14 dans le cas de la séquence sans nuages. d) *Images-résumé* des motifs de longueur 15 dans le cas de la séquence avec nuages. e) *Images-résumé* des motifs de longueur 14 dans le cas de la séquence avec nuages.

Sixième partie

Conclusion et perspectives

10.4 Conclusion

Contrairement à des STIS à résolution grossière pour lesquelles on cherche à séparer un pixel car il contient plusieurs objets d'intérêt, à haute résolution, les pixels doivent être regroupés pour constituer des objets. En effet, à haute résolution, des objets deviennent visibles, et leur forme prend de l'importance. La forme est une des caractéristiques d'un objet, et dans le cas d'une scène dynamique, cette dernière peut aussi évoluer. Afin d'extraire les objets, et d'analyser leurs évolutions, des hypothèses sur les types de changements existants doivent être formulées. Les STIS-HR ne sont pas inter-calibrées radiométriquement. De plus, contrairement à la plupart des séquences d'images existantes, les objets de la scène observée présentent des changements radiométriques dont l'amplitude peut être importante. En conséquence, considérer les objets constants pour inter-calibrer les images comme mentionné ci-dessus n'est pas possible, et a fortiori, l'homogénéité radiométrique ne peut être exploitée pour extraire les objets. En revanche, les images sont supposées recalées à de faibles erreurs près.

Lorsqu'on considère des scènes dynamiques, la troisième dimension, qui peut être ignorée dans le cas statique doit être prise en compte pour expliquer les occultations et désoccultations. Nous avons donc proposé, dans le chapitre 4, un modèle par couches d'objets plats afin d'expliquer les recouvrements observables tels que le recouvrement d'un nuage, d'une traînée de fumée, ou à très haute résolution, de véhicules. L'arrière-plan des scènes de télédétection est très particulier, et son analyse constitue la majeure partie de ce mémoire.

Les objets qui le constituent sont immobiles mais évoluent radiométriquement entraînant des changements topologiques de type fusion, division, et la combinaison des deux. Nous avons proposé une représentation de la dynamique de cet arrière-plan dans un graphe d'adjacence temporelle des objets (SOTAG). Les nœuds de ce graphe représentent les objets de la scène, et les arcs les correspondances temporelles entre ces objets traduisant une relation de type *devient*. Une navigation dans ce graphe permet donc d'accéder à l'histoire des objets.

Dans la deuxième partie du mémoire (ch.7 et 8), nous avons présenté comment extraire ce graphe d'une STIS. Les arcs traduisant des correspondances entre objets et groupes d'objets, il paraît naturel de décomposer la construction du graphe en une étape d'estimation des objets, suivie d'une estimation de leurs correspondances. Cependant, découpler totalement ces deux étapes fournit des graphes très complexes dont les objets ne correspondent pas nécessairement aux objets d'intérêt. Il existe en effet un grand nombre de partitions en zones homogènes possibles pour une image. Nous avons donc proposé un chemin d'optimisation des partitions permettant de prendre en compte la redondance temporelle. Nous exploitons pour cela une heuristique selon laquelle les objets ne changent généralement pas de forme. Dans ces conditions, la forme d'un objet à un certain instant peut être décrite par rapport à la forme de l'objet correspondant dans l'image précédente. Dans les cas où l'objet n'a pas changé, seule la désignation de l'objet correspondant est nécessaire, et plus l'objet diffère, plus la description de la forme est complexe jusqu'au cas extrême où la forme est décrite dans son intégralité. Afin de mesurer les complexités relatives aux différentes descriptions, nous nous sommes placés dans le cadre du MDL, et une longueur de description géométrique spatio-temporelle a été proposée. Cette longueur de description est constituée de deux termes : un terme temporellement localisé décrivant la forme locale des régions, et un terme temporellement dépendant décrivant les modifications à effectuer pour passer d'une forme à l'autre. Les objets de la scène subissent des divisions et des fusions. Cependant, afin de simplifier le problème lors de la construction des partitions,

seules des correspondances 1-1 ont été considérées. La description considérée est donc redondante dans les cas de correspondances 1-n ou n-m, et la longueur de description est alors sur-estimée. Nous avons donc mis à échelle les termes de fidélité et de description du modèle de façon empirique.

Une fois les objet extraits, les nœuds du graphe d'adjacence temporelle des objets sont déterminés, et les arcs peuvent alors être trouvés en estimant les correspondances entre objets. Tous types de correspondances sont considérés : correspondance 1-1 (conservation), correspondance 1-n et n-1, et correspondance n-m. Les correspondances candidates sont les arcs du graphe d'intersection entre les partitions des différentes images. Dans ce graphe, beaucoup d'arcs sont dûs à de petites déformations des régions, liées à de petites déformation des objets, à des erreurs de segmentation ou à des défauts de correction géométrique. Les correspondances n-m sont très difficiles à trouver. Cependant, en remarquant que les objets qui sont en correspondance n-m ne sont ni en correspondance 1-1, ni en correspondance 1-n, nous écartons ce problème en considérant que les correspondances n-m sont les arcs qui restent dans le graphe d'intersection après avoir enlevé tous les arcs de correspondance 1-1, 1-n, et n-1. Afin de ne sélectionner que les arcs traduisant une correspondance 1-1 ou 1-n pertinente des objets de la scène, nous avons adopté une approche de type *winner-take-all* pour un critère de mise en correspondance de plus en plus souple. Le graphe ainsi obtenu permet une navigation vers le futur et le passé d'un certain objet.

Cependant, cette représentation est temporellement localisée dans le sens où les objets d'un certain instant sont liés à leur correspondants dans les images précédentes et suivantes mais pas au-delà. En conséquence, tous les chemins passant par un nœud ne sont pas pertinents pour décrire son histoire, particulièrement dans le cas où elle est impliquée dans une fusion avec un grand nombre de régions puis à nouveau divisée. Afin de permettre une description de l'histoire à long terme d'un objet, l'utilisation jointe de l'intersection des partitions permet une sélection des chemins pertinents.

En considérant une version attribuée du graphe, le chemin correspondant à l'histoire d'un objet contient aussi son évolution radiométrique. Cependant, certaines valeurs de la séquence ainsi obtenue ne correspondent pas à l'objet d'intérêt. Elles peuvent correspondre à un objet le recouvrant, ou à des objets adjacents en raison d'éventuelles erreurs de segmentation ou de mise en correspondance. Nous avons proposé une méthode permettant de regrouper de façon automatique les objets sous-jacents par similarité d'évolution radiométrique malgré l'existence de telles valeurs. Dans ce but des sous-séquences de ces chemins de longueur variable et de support temporel non compact sont regroupées. Un grand nombre de méthodes de recherche de motifs fréquents dans des bases de données linéaires existent. Parmi elles, l'algorithme ECLAT [Zaki, 2000c, Zaki et al., 1997] permet une recherche efficace des motifs maximaux dans le cas où un grand nombre de motifs fréquents existent. Cet algorithme construit l'ensemble des motifs fréquents en parcourant en profondeur l'arbre de préfixe les représentant. Comme la plupart des algorithmes de recherche de motifs fréquents dans des bases de données linéaires, cet algorithme s'appuie sur la propriété d'anti-monotonie de la fréquence. Cette propriété n'est plus vérifiée dans le cas des graphes que nous considérons, une étape de croissance des motifs par fermeture a donc été proposée.

Cette méthode fournit donc de façon automatique les motifs fréquents et capture ainsi la structure présente dans les données. L'interprétation des motifs obtenus n'est néanmoins pas évidente, et une interprétation théorique utilisant des raisonnements a été présentée.

Nous avons donc présenté une méthode de segmentation spatio-temporelle d'une STIS à haute résolution qui consiste dans un premier temps en un partitionnement spatial de

chaque image conditionnellement à la STIS, et dans un second temps en une sélection temporelle d'instantants pertinents.

10.5 Perspectives

Le modèle de scène que nous avons considéré est constitué de plusieurs couches, et nous nous sommes uniquement intéressés dans ce mémoire à l'arrière-plan. La prise en compte des autres couches permettrait d'éviter l'apparition de valeurs incohérentes dans les séquences qui compliquent l'analyse radiométrique. En effet, dans ces valeurs, une estimation de motifs temporellement compacts serait alors possible.

De plus, dans le modèle que nous avons proposé, les objets des autres plans sont opaques et radiométriquement constants. Or les nuages, les zones d'humidité ou encore les zones d'ombre ne sont pas toujours opaques, et il pourrait être intéressant d'exploiter l'information temporelle pour permettre une estimation de l'arrière-plan qui se trouve dessous dans les cas où ils ne sont pas opaques.

Nous avons montré que l'énergie considérée pour la segmentation permet effectivement de capturer l'information de redondance de forme des objets. Cependant, l'algorithme d'optimisation devient très lent lorsque la taille des images de la séquence croît. En début d'optimisation, les fusions entre régions suffisamment éloignées sont indépendantes. Nous avons donc proposé une parallélisation en début d'algorithme de plusieurs optimisations sur des zones de taille réduites découpées dans la STIS. Ce protocole nous a permis d'obtenir des résultats de segmentation sur des images plus grandes. Cette approche peut présenter des défauts liés au découpage initial arbitraire, et une amélioration de l'optimisation reste nécessaire. Dans ce but, on pourra considérer un algorithme multi-échelles permettant d'effectuer les optimisations indépendantes sur des zones correspondant à des réunions de régions obtenues par une multi-fusion sur un sous-échantillonnage spatial de la STIS. Notons ici que ce genre d'approche est plus compliqué qu'en 2D puisque les régions peuvent ne pas se correspondre entre les différentes images.

Pour la construction du graphe, aucune sélection des arcs n'est effectuée pour les correspondances n-m qui sont simplement *celles qui restent* lorsqu'on a enlevé les correspondances 1-1 et les correspondances 1-n. En conséquence les correspondances n-m peuvent être très complexes. Intégrer dans l'algorithme proposé une estimation de ces correspondances permettrait une simplification supplémentaire du graphe.

Par ailleurs, le graphe considéré est un graphe d'adjacence transitif. En conséquence, obtenir l'histoire à long terme d'un objet nécessite une sélection des chemins passant par le nœud d'intérêt. Nous avons effectué cette sélection en utilisant une intersection entre toutes les images de la séquence. Une perspective intéressante serait de trouver un moyen d'intégrer la connaissance des chemins d'intérêt directement dans la structure.

Lors de la recherche de motifs fréquents dans le graphe d'adjacence temporelle des objets, nous avons effectué une croissance des motifs par fermeture. En conséquence, certains motifs fréquents retrouvés par l'algorithme peuvent avoir un motif couvrant fréquent que l'algorithme n'a pas retrouvé. Par exemple, supposons que DE soit un motif fréquent maximal de l'arbre de préfixe. Supposons de plus que $ACDE$ et que $BCDE$ soient deux motifs fréquents, A et B correspondant à des nœuds communs à un grand nombre de chemins du graphe. Notre algorithme ne trouvera que CDE . Une croissance des motifs par construction d'autres arbres de préfixe conditionnels aux motifs fréquents permettrait d'obtenir des motifs plus longs. Notons que dans ce cas, la complexité est sensiblement

acruë car la structure alors considérée est un premier arbre dont chaque feuille est la racine d'un autre arbre de profondeur réduite.

Nous avons proposé une méthode automatique pour la classification des évolutions. L'interprétation des motifs fréquents obtenus par cette étape est délicate. Cependant, l'information ainsi extraite est pertinente car elle est fréquente. La question de savoir comment l'exploiter reste encore grand ouverte. L'utilisation de raisonnements s'appuyant sur des ontologies est une réponse possible.

Une alternative à l'approche automatique est une approche supervisée au cours de laquelle un utilisateur fournirait des exemples d'objets d'une même classe.

L'analyse de la STIS que nous proposons se décompose en trois étapes : la segmentation des images, la construction du graphe, et un regroupement des évolutions radiométriques similaires. Nous avons présenté une méthode de multi-segmentation jointe permettant de ne pas découpler totalement la première et la deuxième étape. Coupler de la même façon l'étape deux et l'étape trois permettrait d'éviter les arcs liés à des erreurs d'appariements qui viennent perturber l'analyse radiométrique.

Annexe A

Annexes

A.1 Le principe de minimisation de longueur de message (MDL)

Le MDL formalise le principe du rasoir d'Ockham selon lequel le meilleur modèle est le plus simple. Il a été présenté pour la première fois dans [Rissanen, 1983] pour estimer un modèle de données, son ordre étant inconnu. Comme Kolmogorov a montré qu'il était impossible de trouver le plus petit programme permettant de générer des données (ce qui correspondrait à la description la plus courte), Rissanen propose de ne pas répondre à cette question de façon absolue, mais relative, en considérant un ensemble de modèles, \mathcal{M} pour les données, D , et en sélectionnant celui qui permet de décrire les données de la façon la plus courte.

$$\hat{\mathcal{M}} = \operatorname{argmin}_{\mathcal{M}} (\mathcal{L}(D|\mathcal{M}) + \mathcal{L}(\mathcal{M})) \quad (\text{A.1})$$

La longueur de description est définie par la longueur de code optimale de Shannon : $\mathcal{L}(x) = -\log(P(x))$. Dans le cas où l'ensemble des modèles est discret et fini, cette expression est alors équivalente au maximum a posteriori.

Cependant, seules des données discrètes auront une telle longueur de code correspondante finie. Or, en réalité, l'ensemble des modèles n'est ni discret, ni fini : les paramètres des modèles qu'on utilise sont généralement continus. Rissanen propose donc une méthode de troncation des paramètres des modèles permettant de calculer cette expression. La précision optimale est déduite de la minimisation de la longueur de description pour la pire valeur de paramètre trouvée dans le rectangle associée à cette précision.

Wallace propose dans [Wallace and Freeman, 1992] une alternative à cette estimation qui consiste à considérer que le paramètre réel est uniformément distribué dans un intervalle de longueur la précision. Il optimise ensuite la précision en considérant le cas moyen. La longueur de message qui en découle s'appelle "Minimum Message Length" (MML) pour se différencier de la "Minimum Description Length" (MDL) de Rissanen.

L'approche de Rissanen a l'avantage de n'effectuer aucune hypothèse sur la distribution du paramètre. De façon générale, la volonté de trouver un modèle en considérant qu'on n'a aucun *a priori* sur les données est ce qui différencie l'approche de Rissanen de celle de Wallace. Notons cependant pour éviter des *qui pro quo* que si Rissanen ne suppose aucun prior sur la distribution du paramètre réel dans son rectangle de précision, il suppose en revanche une répartition uniforme du paramètre tronqué [Lantermann, 2001]. Rissanen a de plus défini un "codage universel pour les entiers" qu'il dérive de la minimisation de la longueur de code moyenne d'une distribution décrivant la totalité des données (modèle

et résidus) dans le cas du pire modèle. La longueur optimale est finalement : $\mathcal{L}(x) = \log^*(x) + \log(c)$ où $c \simeq 2.865064$.

Un autre avantage de cette approche est qu'elle est sans paramètre libre contrairement à d'autres méthodes de régularisation où le compromis entre la simplicité et la fidélité est réglé par l'utilisateur. Enfin, la mixité de la représentation (sous forme déterministe en comptant le nombre de bits ou à partir d'un modèle probabiliste en utilisant la longueur de message de Shannon) est un avantage non négligeable pour répondre au problème de la segmentation.

Les modèles considérés sont liés aux propriétés des données car ils exploitent leurs redondances pour définir un espace de représentation adapté. Ces redondances diffèrent en fonction de la scène imagée, et des phénomènes d'intérêt. Notre choix pour une approche sera donc influencé par les redondances qu'elle exploite qui doivent correspondre aux redondances de nos données. Dans la littérature, le modèle des données n'est pas toujours présenté de façon explicite, mais peut être implicite à une méthode d'analyse ou de compression. Nous proposons donc un état de l'art des méthodes d'analyse et de compression de séquences d'images mettant l'accent sur le modèle considéré bien que l'explicitation de ce dernier soit souvent omise dans la littérature en raison de son évidence pour la communauté du domaine.

Dans ce chapitre, nous présentons différents modèles de séquences d'images existants dans la littérature. Cette annexe est le pendant de la section ?? dans laquelle ces modèles ont été passés en revue de façon succincte de façon à positioner notre problème par rapport à l'existant.

Nous nous intéresserons dans un premier temps à des séquences de fluides. Les séquences de fluides présentent, comme les STIS-HR une variabilité radiométrique temporelle. Nous verrons succinctement les approches adoptées en imagerie médicale pour étudier le mouvement d'organes qui analyse des déformations. Puis, nous verrons les approches utilisées en multimédia pour lesquelles un mouvement rigide est généralement supposé.

A.2 Etat de l'art de l'analyse de séquences d'images

A.2.1 Séquences d'images de fluides

Les fluides sont déformables, et la "forme" d'un fluide a peu de sens. Ainsi, les entités d'intérêt seront les pixels, et non pas des régions. Deux types de modèles sont généralement considérés : ceux qui supposent la conservation de l'intensité des pixels [Horn and Schunck, 1981, Suter, 1994, Gupta and Prince, 1996], et ceux qui autorisent une variation d'intensité. [Corpetti et al., 2000] considère ainsi une évolution radiométrique des pixels au cours du temps. L'équation de contrainte d'image, ¹ est alors remplacée par une *équation de continuité* de façon à prendre en compte des variations de l'intensité le long des déplacements. Cette équation provient de la loi de conservation de la matière (la variation du volume correspond à un flux normal à sa surface) et contient donc un terme supplémentaire : la divergence de la vitesse multipliée par l'intensité traduisant le fait que les zones plus denses auront des vitesses plus rapides de façon à uniformiser la densité. Dans ce cas, de l'information *a priori* sur l'évolution des pixels issue d'un modèle physique est utilisée pour analyser l'évolution de la radiométrie le long des déplacements.

Dans le cas de l'analyse des mouvements atmosphériques par imagerie satellitaire, l'image acquise est une image 2D alors que la scène est tridimensionnelle. Or l'équation

1. aussi appelée équation du flux optique.

de continuité appliquée à l'image suppose que les transferts de matière sont horizontaux. Ainsi, deux nuages d'altitude différente (donc n'entrant pas en interaction dans la scène) interagissent dans le modèle 2D considéré appliqué à l'image. [Heas et al., 2007] propose un modèle multi-couches permettant de prendre en compte la troisième dimension et de considérer que seuls des nuages de même altitude peuvent interagir.

A.2.2 Séquences d'images médicales

Une difficulté que présente l'analyse de la STIS est que les objets évoluent en forme et radiométriquement et de plus que les images sont mal inter-calibrées radiométriquement. On retrouve ces difficultés dans une application de l'analyse d'images médicales lors du suivi de patients atteints de sclérose en plaques. En effet, la forme et la radiométrie des lésions évoluent au cours du temps et les images ne sont pas intercalibrées entre les différentes acquisitions. [Gerig et al., 2000, Bosc et al., 2003] proposent des calibrations non linéaires des radiométries s'appuyant respectivement sur une modélisation physique de l'acquisition et sur des histogrammes joints. Les changements de formes considérés sont des déformations et des erreurs de recalage existent généralement entre les images. L'estimation de la déformation est généralement effectuée après l'estimation des déplacements rigides liés à un décalage entre les images. Différentes méthodes d'estimation de mouvement non rigide existent, par exemple [Thirion, 1998], qui propose une estimation du flux optique s'appuyant sur un modèle de diffusion de telle sorte que les points des différentes images s'attirent et se repoussent selon leur différence radiométrique. Une méthode plus robuste aux erreurs d'intercalibration utilisant des mises en correspondances de statistiques d'ordre supérieur est proposée dans [Cachier and Pennec, 2000].

De façon générale, ces méthodes gèrent le problème d'évolution d'objets en forme et en radiométrie conjointement à des erreurs d'intercalibration géométrique et radiométriques par des méthodes séquentielles supposant les unes constantes pour permettre l'estimation des autres. Ainsi, dans [Rey et al., 2002, Bosc et al., 2003], les images sont d'abord supposées radiométriquement calibrées et les objets immobiles pour permettre un recalage rigide des images. Sur les images ainsi recalées, on suppose ensuite qu'il n'y a que des déformations non rigides, puis après avoir estimé ces dernières, [Bosc et al., 2003] suppose les images mal inter-calibrées mais les objets constants, puis après estimation de l'inter-calibration, le fond est supposé constant et les objets varient radiométriquement de sorte qu'une détection de changement est possible.

D'autres séquences classiquement analysées en imagerie médicale sont les séquences d'images cardiaques. En effet, le coeur est un organe dont l'étude du mouvement présente un intérêt clinique et une littérature abonde donc sur le sujet. Une première catégorie d'analyse qui s'est développée dans les années 80 utilise des *a priori* forts sur les types de mouvements effectués pour compenser le peu d'images temporelles et de coupes. Des paramètres du mouvement présentant un intérêt clinique sont alors extraits des images [Ingels et al., 1978, Leighton et al., 1974, Stewart et al., 1975, Daughters et al., 1982, Bolson et al., 1980]. A partir des années 90, les progrès techniques dans l'acquisition permettent de considérer une modélisation 3D+t de l'évolution de la paroi. Dans [Duncan et al., 1994, Kambhamettu and Goldgof, 1994], une mise en correspondance point à point entre les parois des deux images est effectuée de façon à modéliser la déformation. Le problème d'estimation d'une déformation 3D d'un volume à partir de coupes est cependant sous-déterminé, et des contraintes de lissage doivent donc être considérées. Dans [Park et al., 1996a,b] un système d'imagerie par résonance magnétique (IRM) particulier per-

met d'obtenir une mise en correspondance de points organisés sur une grille. [Matheny and Goldgof, 1995] propose une extension des paramètres de forme que sont les harmoniques sphériques à la dimension temporelle. Dans la même lignée, [Clarysse et al., 1997] propose aussi une analyse temporelle des paramètres de forme pour une séquence d'images 3D IRM en construisant le spectre de caractéristiques de courbures globales, et locales.

A.2.3 Séquences vidéo multimedia

On peut distinguer deux types de partitionnement de vidéo. La segmentation temporelle consiste à segmenter une séquence d'images en un ensemble de sous-séquences temporelles d'images caractérisées. Cette étape revient donc à segmenter l'axe temporel. Les sous-séquences obtenues sont appelées *plans*, et présentent une unité de contenu de sorte qu'un second type de partitionnement, spatio-temporel, de ces sous-séquences peut par la suite être effectué.

A.2.3.1 Segmentation temporelle

Un plan correspond à un événement qui se produit dans un certain fond, tandis que la scène est une juxtaposition de plans dont la logique permet d'exprimer une idée plus élaborée. En conséquence, les vidéos sont segmentées temporellement en plans pour conduire l'analyse et l'extraction d'information. Ces méthodes correspondent donc à des changements à la fois structurels et radiométriques. Etant donné que les STIS présentent aussi des changements structurels, nous présentons succinctement les approches généralement employées dans ce domaine et pour une revue complète de ces méthodes, le lecteur est renvoyé à [Koprinska and Carrato, 2001].

Remarquons d'abord que si les changements sont structurels et radiométriques, ils ne sont cependant pas nécessairement brutaux. En effet, des effets spéciaux peuvent être utilisés entre différentes scènes amenant à différents types de ruptures de plan : le fondu entrant, fondu sortant, balayage, dissolution, et coupure.

La compression de vidéo est parfois précédée d'une étape de détection de (changement de) plans de façon à regrouper les images par plans puis par scène, et les compresser indépendamment [Mech and Wollborn, 1998]. En effet, compenser des images en mouvement est simple et avantageux (d'un point de vue compression) pour des images d'un même plan. En revanche, entre deux plans, il serait plus intéressant de coder la première image du nouveau plan plutôt que de coder une image différente.

Il existe différentes méthodes de détection de transitions. La plus simple consiste à calculer, pour chaque paire d'images successives une mesure de dissimilarité radiométrique. Cette dissimilarité est ensuite seuillée de façon à détecter les transitions. Trois mesures de dissimilarité sont proposées dans [Zhang et al., 1993] : le nombre de pixels changés obtenus par seuillage d'une image de différence (éventuellement compensée en mouvement), le nombre de blocs changés, la décision de changement étant prise selon un critère de vraisemblance, et une distance entre histogrammes. L'auteur propose de plus un seuil adaptatif afin de prendre en compte des scènes plus ou moins mobiles. [Hampapur et al., 1995] propose de gérer les transitions de fondu en effectuant une *mise à échelle chromatique* : plutôt qu'une simple différence d'images, la différence divisée par l'intensité dans l'image précédente est utilisée comme mesure de dissimilarité. [Zabih et al., 1995] propose une mesure de dissimilarité basée sur le taux d'apparitions/disparitions de contours de façon à être insensible aux effets spéciaux liés à la radiométrie.

[Boreczky and Wilcox, 1998] propose d'utiliser des champs de Markov pour modéliser les différents états de la vidéo : plan, coupure, fondu, balayage, déplacement de caméra, et zoom. [Hanjalic et al., 1999] propose de regrouper ces plans une fois identifiés par des similarités de contenu de plus haut niveau de façon à mettre en évidence des scènes.

A.2.3.2 Analyse du mouvement dans les séquences multimédia par estimation du flux optique

Le flux optique, par définition, est le mouvement apparent sur les images. C'est donc une estimation de la vitesse d'un objet dans la scène. Le flux optique peut être vu comme une manifestation d'une force de déformation sur les images. Cette approche est devenue une méthodologie classique pour l'estimation du mouvement. La sous-détermination de ce problème amène à effectuer des approximations dont l'une est en particulier que le mouvement est spatialement lisse. Ce modèle colle assez peu à la réalité dans le cas de contenu multimédia car les séquences sont constituées d'objets aux frontières franches dont les mouvements peuvent différer beaucoup. Un modèle par morceaux est donc plus réaliste dans le cas de contenu multimédia. Cependant, l'estimation du mouvement est un problème qui se mord la queue car pour estimer le mouvement, il faut connaître des mises en correspondances, et pour connaître des mises en correspondances, il faut connaître le mouvement.

A.2.3.3 Segmentation spatio-temporelle de vidéos

Une fois les plans extraits, on peut, dans un but de compression ou d'analyse s'intéresser aux différents objets de la scène. Cette étape est appelée *segmentation spatio-temporelle*. Il existe trois approches pour cette étape : l'une consiste à segmenter spatialement les images de la séquence selon une homogénéité de mouvement, l'autre consiste à regrouper temporellement des régions extraites dans chaque image de façon à mettre en évidence les objets spatio-temporels de façon explicite, enfin, un dernier type d'approche utilisé pour le codage à très bas coût est le codage basé modèle. Cette dernière approche considère un modèle d'objet dont les différentes instances temporelles sont obtenues par déformation.

Segmentation spatiale en mouvement La segmentation en mouvement consiste à partitionner les images en régions homogènes en mouvement. Cette information permet en effet de générer une image à partir de la précédente ce qui présente un intérêt du point de vue du codage. Cette image générée à partir de l'information de mouvement est appelée *compensée en mouvement*. Par ailleurs, les régions homogènes en mouvement sont généralement plus proches d'objets sémantiques et présentent donc un intérêt particulier aussi d'un point de vue analyse.

[Diehl, 1991] propose une méthode de segmentation en mouvement et en profondeur. Un mouvement rigide des objets est considéré et une hypothèse d'invariance radiométrique des objets est implicitement considérée pour l'estimation de ce mouvement. L'estimation du mouvement des objets et des objets eux-mêmes est effectuée simultanément dans un algorithme de segmentation hiérarchique extrayant successivement les composantes connexes dont le mouvement est proche du mouvement dominant.

Dans [Gelgon and Bouthemy, 2000], les auteurs considèrent un regroupement de régions spatiales par la similarité du mouvement qu'elles présentent. La similarité spatiale de mouvement basée sur l'erreur de compensation est utilisée pour modéliser l'interaction

entre régions dans un champ de Markov. Le même type de dissimilarité est considéré dans [Wang, 1998].

Parmi les approches qui exploitent la similarité en mouvement, on trouve les travaux de [Wang and Adelson, 1994] faisant intervenir une modélisation de la scène par plusieurs couches ayant un mouvement propre. Dans cet article, le modèle est présenté de façon assez explicite : la scène considérée consiste en plusieurs couches d'extensions spatiales variables, et ordonnées selon leur profondeur. Une segmentation en mouvement suivie d'un cumul des différentes couches compensées en mouvement permet d'obtenir une pyramide de la scène. Ce modèle de couches a aussi été utilisé dans [Ke and Kanade, 2001], et [Xiao and Shah, 2005]. Notons que la représentation de la scène sous forme d'une image n'est possible que si la caméra a un mouvement de translation. En effet, dans le cas où la caméra tourne autour d'un objet, il est impossible de représenter les différentes vues par une simple image : le résultat obtenu dans ce cas ressemble à une projection cylindrique de l'objet. De plus, l'hypothèse de mouvement rigide contraint fortement le type de scène descriptible. Ce type d'approche peut donc être utilisé pour modéliser un fond, mais d'autres approches doivent être appliquées pour décrire le mouvement d'objets de premiers plans.

Les régions obtenues par similarité du mouvement présentent souvent des imprécisions au niveau des contours liés aux méthodes d'estimation du mouvement. Cependant, ces régions correspondent à des réunions d'objets homogènes en radiométrie. En conséquence, leurs contours doivent coïncider. Afin de corriger les imprécisions des contours, des méthodes utilisant conjointement les deux types d'information ont été proposées. Ainsi, [De Roover et al., 2005] intègre l'information spatiale en considérant un potentiel supplémentaire obtenu à partir d'une segmentation par LPE dans l'équation de mouvement d'un contour actif.

Segmentation spatio-temporelle en objets spatiaux temporellement cohérents

Les objets précédemment extraits ne sont pas explicitement liés temporellement. Afin d'analyser les trajectoires d'objets, ou pour permettre des fonctionnalités basées contenu, et de façon générale pour caractériser un objet spatio-temporel dans son intégralité, les objets spatiaux –quelle que soit leur nature– doivent être mis en correspondance.

Dans ce but, [Marques and Molina, 1997, Marcotegui, 1996] proposent de fusionner temporellement des régions spatiales connexes au sens de la connexité d'un espace 3D. Afin de permettre la connexité temporelle des différentes instances des objets, une compensée en mouvement des partitions spatiales est utilisée. Des marqueurs sont propagés selon des informations de différents types (contraste, texture et mouvement) servant d'initialisation à un algorithme de LPE. La redondance temporelle exploitée pour les extraire est une redondance de mouvement rigide, et le raffinement par LPE après projection autorise une légère déformation des objets. L'estimation du mouvement suppose par ailleurs une invariance radiométrique temporelle des objets. Afin d'extraire des objets à caractéristiques sémantiques, [Marcotegui, 1996] propose de plus d'introduire ce type d'information de façon interactive à partir d'un trait dessiné sur une image de la séquence par l'utilisateur. Dans la même lignée, [Gu and Lee, 1997] propose une méthode de segmentation sémantique suivie d'une méthode de poursuite basée sur la morphologie et un modèle de mouvement tridimensionnel. [Reusens, 1995, Salembier et al., 1997] considèrent des objets de natures différentes codés par des schémas de codage différents de façon à optimiser le compromis débit-distorsion. Afin de faciliter ce choix, une hiérarchie de segmentations combinant des fusions par mouvement ou par radiométrie est propagée. [Bouthemy and Francois, 1993] présentent une méthode de segmentation en mouvement temporellement cohérente utilisant

des champs de Markov dont les observations sont les dérivées spatio-temporelles de la séquence en tout point. Dans [Moscheni et al., 1998], les régions sont fusionnées selon un critère spatio-temporel utilisant comme homogénéité temporelle le mouvement, et comme homogénéité spatiale la radiométrie. Les régions sont ensuite fusionnées en même temps dans le temps et dans l'espace via un algorithme de fusion de régions utilisant un graphe d'adjacence spatio-temporelle de régions, l'adjacence temporelle liant des régions d'images consécutives.

Une autre méthode, appelée *projection de partitions* est utilisée pour le codage de vidéo. Cette méthode s'appuie sur le concept d'images-I, et d'images-P. Il s'agit respectivement d'images codées en intra (en ne considérant aucune autre image), et par rapport aux précédentes. En effet, les paramètres de mouvement de région peuvent être assimilés aux paramètres de compensation de texture. Cependant, cette approche gère mal les occlusions. [Marques et al., 1998, Bremond and Marquès, 1996] proposent donc de coder séparément les partitions, leurs paramètres d'évolution, et la texture et ses paramètres. [Bremond and Marquès, 1996] propose dans ce but une méthode d'interpolation et de prédiction de partition. Des paramètres d'évolution sont calculés pour chaque région de façon à estimer les régions interpolées correspondantes. Les régions sont ensuite déplacées et déformées successivement par profondeur décroissante grâce à une déformation de partition utilisant une distance géodésique s'appuyant sur la distance de Hausdorff. Cependant, la distance géodésique se calculant entre deux images, seuls les cas d'interpolation (non causale) sont pris en compte. [Marques et al., 1998] propose donc d'utiliser des descripteurs de Fourier afin de prendre en compte aussi les cas d'extrapolation.

Enfin, une fois ces séquences d'objets spatiaux (OS) identifiées, une estimation d'un mouvement temporellement cohérent sur des longues séquences peut alors être effectuée. Ainsi, [Dufaux and Moscheni, 1996], proposent une estimation du mouvement des régions obtenues par segmentation spatio-temporelle le long de la séquence par filtrage Kalman. Dans [Meyer and Bouthemy, 1994, Bouthemy and Francois, 1993] les objets issus de la segmentation spatio-temporelle cohérente en mouvement décrite dans [Bouthemy and Francois, 1993] sont poursuivis en supposant des trajectoires bidimensionnelles dans le plan de l'image. Des hypothèses de continuité du mouvement sont aussi exploitées pour faire face au problème d'occlusion en rejoignant les différentes portions de trajectoires.

Comme nous l'avons mentionné, le problème de l'estimation du mouvement est un problème qui se mord la queue. Une première estimation doit être effectuée afin d'obtenir des régions sur lesquelles calculer des paramètres de mouvements affines. Nous avons jusqu'à présent vu des méthodes extrayant l'information de mouvement à partir d'un calcul de flux optique par blocs. Une autre possibilité est de considérer des cartes de changements [Ohm, 1998, Meier and Ngan, 1998, Mech and Wollborn, 1998, Neri et al., 1998, Guo et al., 1999]. Une carte de changement est obtenue à partir d'un seuillage sur la différence d'images. [Ohm, 1998] considère des statistiques d'ordre supérieur (moment d'ordre 4) afin de tenir compte de la non normalité de l'image de différence. Ces approches mettent en évidence des changements au niveau des frontières. Des outils morphologiques sont donc souvent utilisés pour combler les trous. [Meier and Ngan, 1998] considèrent de plus que la radiométrie des objets peut varier. Pour ce faire, la redondance temporelle des objets exploitée n'est plus la radiométrie, mais les contours extérieurs *et* intérieurs des objets mobiles. Ces objets mobiles, sont obtenus à partir de cartes de changements appliquées aux images de gradient (spatial). Puis, les objets sont mis en correspondance en minimisant une distance de Hausdorff entre les différents contours des objets mobiles. Les nouveaux objets sont mis à jour dans chaque image de façon à prendre en compte des déformations,

rotations et translations. Cette approche présente l'avantage de s'affranchir de la contrainte d'invariance radiométrique des objets.

Codage basé modèle Ces méthodes sont des approches de codage qui consistent à synthétiser les images de la séquence à partir des paramètres d'un modèle. Le modèle considéré comprend généralement différents objets tri-dimensionnels, qui se déforment au cours du temps. Ainsi, les paramètres de forme, et de texture de chacun des objets ne sont transmis qu'une fois, et, seuls les paramètres de déformation (aussi appelés paramètres d'animation) ont à être transmis par la suite. Ces approches qui permettent d'atteindre de très forts taux de compression sont apparues dans le contexte de la visioconférence pour lequel le débit est la principale contrainte. La grande limitation de ces types d'approches est que le modèle des objets doit être connu *a priori*. Le type de scène que l'on peut traiter par cette méthode est donc assez restreint, et les applications se limitent généralement à de la visiophonie et de la transmission télévisée pour laquelle, seules tête et épaules d'une personne doivent être représentés. [Eisert and Girod, 1998] propose l'utilisation de B-splines pour interpoler le visage à partir des points de contrôle. Les expressions du visage sont modélisées par une superposition d'actions unitaires sur le modèle.

L'approche par synthèse est aussi celle adoptée par le standard MPEG4 qui suppose un modèle par couches permettant de transmettre une seule fois le fond (après une segmentation temporelle en plans d'objets vidéo, VOP), sur lequel des objets transmis de façon indépendante (la texture étant éventuellement aussi transmise qu'une seule fois) sont incrustés [ISO/IEC, 1998].

Conclusion Nous avons vu que les méthodes d'analyse de vidéo classiquement employées permettent d'extraire deux types de modèles de scènes spatio-temporelles : des scènes constituées d'objets temporels et des scènes constituées d'un ensemble d'objets spatio-temporels. Les objets temporels sont des ensembles d'objets spatiaux de vitesse similaire, et les objets spatio-temporels sont des séquences d'objets spatiaux. Ces objets spatiaux peuvent présenter une homogénéité radiométrique, de mouvement, de profondeur, sémantique et les séquences d'objets spatiaux peuvent être caractérisées par une redondance temporelle de mouvement, de radiométrie, de forme ou encore une combinaison de ces caractéristiques.

Deux types de mouvements sont considérés pour les objets : les mouvements rigides et les mouvements de déformation. L'estimation de mouvement est généralement basée sur une homogénéité temporelle de mouvement par une optimisation itérative. Mais elle peut aussi s'appuyer sur une homogénéité radiométrique, de forme, de points caractéristiques, ou encore combiner ces différentes approches. De petites déformations peuvent être prises en compte de façon implicite en considérant des homogénéités radiométriques temporellement localisées, et des déformations plus importantes peuvent être prises en compte avec des méthodes basées modèle.

Le tableau 1.1 synthétise l'état de l'art sus-détaillé en mettant l'accent sur le type de modèle appliqué aux données.

A.2.4 Mouvement dans les images satellitaires de la terre

Des séquences à basse résolution spatiale mais à haute fréquence temporelle permettant de suivre des évolutions de la terre peuvent être acquises. Ainsi, les satellites de

NOAA portent un radiomètre à très haute résolution (radiométrique), AVHRR permettant d'acquérir des données de résolution spatiale 1.1km dans cinq canaux allant de la fin du spectre visible à l'infra-rouge, et permet une revisite deux fois par jour. Le spectro-radiomètre imageur à résolution modérée, MODIS, embarqué sur TERRA et sur AQUA permet une fréquence de revisite journalière ou tous les deux jours ; il contient trente-six bandes spectrales dont sept dans le visible, et six dans le proche infra-rouge. Il acquiert des données à résolution spatiale 250m, 500m et 1km (dans l'infra-rouge). Enfin, un capteur d'échelle régionale VEGETATION (VGT) a été embarqué sur SPOT4 (depuis 1998) et SPOT5 (depuis 2003) pour permettre une revisite journalière sur une zone de 2200 km de largeur à résolution spatiale de 1km. Ces capteurs sont caractérisés par quatre bandes spectrales : bleu, rouge, proche infra-rouge et moyen infra-rouge.

De plus, dernièrement, des séquences à haute résolution ont été acquises, et des méthodes d'analyses et de compression ont donc été récemment proposées.

Enfin, il existe un autre champ d'investigation ou des séquences d'au moins deux images satellitaires sont analysées : la détection de changement. De nombreuses méthodes ont été développées dans ce cadre aussi, et nous allons les présenter succinctement bien que l'objectif soit assez différent du nôtre.

A.2.4.1 Analyse de séquence temporelle de la terre à résolution spatiale grossière (STIS-CR)

L'intérêt porté jusqu'à présent sur les évolutions radiométriques des STIS-CR concerne les évolutions de la végétation. Les applications sont diverses, depuis la surveillance de culture pour l'assistance aux agriculteurs, aux études de changements institutionnels sur les pratiques agricoles en passant par l'effet des changements d'occupation du sol pour l'équilibre CO₂-biome-atmosphère qui joue un rôle capital dans le réchauffement climatique.

On peut distinguer deux classes d'applications auxquelles correspondent aussi deux classes de méthodes. Dans un cas, l'objectif est d'obtenir des paramètres phénologiques afin d'évaluer de façon quantitative l'évolution du biome de zones particulières (phénorégions) déterminées indépendamment (généralement à partir de cartes), et dans l'autre cas, l'objectif est de dresser une carte d'occupation du sol. Dans le premier cas, un signal temporel extrait d'une région spatio-temporelle est analysé alors que dans l'autre cas, les signaux temporels correspondant à chaque pixel sont regroupés pour former des régions d'occupation du sol.

Étude phénologique du biome Les études phénologiques s'appuient généralement sur un indice de végétation. Il en existe différents : l'NDVI (NDVI, ²), le simple rapport [Jordan, 1969], l'indice de végétation ajusté sol [Huete, 1988, Qi et al., 1994], l'indice de végétation atmosphériquement résistant [Kaufman and Tanre, 1992], et l'indice de végétation amélioré (EVI, ³) [Huete et al., 1999].

Le NDVI et l'EVI sont maintenant les plus couramment utilisés. Le NDVI exploite deux canaux dans lesquels la végétation a une réponse plus importante : le proche infra-rouge, et le rouge. Il a été relié à différentes grandeurs physiques telles que l'index d'aires de feuilles, ⁴ (LAI) [Spanner et al., 1990], la production de matière sèche au Sahel [Tucker et al., 1983], la production nette primaire en Amérique du nord et du sud [Goward et al., 1985], et [Malingreau, 1986] a montré le lien entre la courbe de NDVI et les dynamiques de

2. Normalized Vegetation Index. Le lecteur est renvoyé au §3.1.5 pour plus de détails sur cet indice

3. Enhanced Vegetation Index

4. L'index d'aires de feuilles est la surface de feuille projetée visible par unité de surface

l'écosystème et des pratiques de culture en Asie. D'autres indices associés à d'autres grandeurs biophysiques existent néanmoins, ainsi, le modèle de Li-Strahler permet d'estimer la couronne des troncs [Macomber and Woodcock, 1994], et d'autres paramètres biophysiques sont proposés dans [LU et al., 2002].

L'EVI exploite de plus la bande bleue afin de réduire l'influence de l'atmosphère ou de fumées. En contrepartie, contrairement au NDVI, il est amplifié par la neige.

Les valeurs des pixels correspondent en effet à l'accumulation de contributions de différentes natures en raison de la résolution spatiale grossière. Ces contributions autres que la végétation d'intérêt sont généralement considérées comme un bruit, et différentes approches de débruitage ont été proposées.

Il existe différentes méthodes de filtrage des séries temporelles de NDVI. [Holben, 1986] introduit la méthode de construction d'images composites de valeurs maximales (MVC). Une telle image composite est obtenue en sélectionnant pour chaque pixel, dans un voisinage spatio-temporel, la valeur de NDVI maximale sur une séquence d'une durée choisie (généralement semaine ou mois). En effet, les effets atmosphériques ont toujours pour effet de réduire la valeur de NDVI, et les effets directionnels ont généralement pour effet de les réduire. Cette méthode présente l'avantage d'être simple. Cependant, les images obtenues peuvent encore être corrompues par la couverture nuageuse, et accumulent les erreurs dites *positives*,⁵. Ce pré-traitement est souvent effectué et nous appellerons les STIS résultantes STIS composites (STISc).

Des méthodes paramétriques ont ensuite été proposées. Elles sont de deux types : les méthodes par ajustement de courbes [Van Dijk et al., 1987, Verhoef et al., 1996, Zhang et al., 2003, Swets et al., 1999] et les méthodes par analyse dans l'espace de Fourier [Olsson and Eklundh, 1994, Moody and Johnson, 2001, Jing et al., 2005]. Dans [Moody and Johnson, 2001, Jing et al., 2005], afin de réduire l'effet d'aliasing mis en évidence dans [Olsson and Eklundh, 1994], trois fréquences seulement ont été sélectionnées : l'harmonique zéro correspondant à la valeur moyenne, la première harmonique correspondant à une période de un an, et la deuxième harmonique correspondant à une période de six mois. Dans, [Moody and Johnson, 2001], les phases et amplitudes correspondant à ces différentes fréquences sont interprétées en termes de caractéristiques phénologiques. Dans [Jing et al., 2005], ces fréquences sont sélectionnées de façon à filtrer le signal. Le filtrage obtenu est cependant généralement trop fort, et restitue mal l'asymétrie des profils phénologiques [Hird and McDermid, 2009, Chen et al., 2004]. [Van Dijk et al., 1987, Verhoef et al., 1996] proposent un lissage par courbes polynomiales. Cependant, ces méthodes lissent généralement trop les données et restituent mal la période de dormance. D'autres méthodes permettant de prendre en compte la nature dissymétrique ainsi que les plateaux des périodes de dormance ont alors été proposées : une méthode d'ajustement par deux fonctions logistiques (DL) [Jonsson and Eklundh, 2002] et une méthode d'ajustement par gaussienne asymétrique [Beck et al., 2006]. On peut cependant reprocher à ces méthodes de sous-estimer la durée de maturation et de ne prendre en compte qu'un cas monomodal de végétation. [Zhang et al., 2003] propose donc une méthode utilisant des séries de fonctions logistiques. L'auteur espère ainsi obtenir une meilleure estimation des paramètres d'intérêt phénologiques que sont la date de début de croissance, le rythme maximal de croissance, les dates de début et de fin de maturité, et la date de début de dormance. De façon générale, les approches

5. Les erreurs négatives sont les erreurs liées à l'acquisition dont l'effet est de réduire le NDVI. Réciproquement, les erreurs positives sont les erreurs dont l'effet est une augmentation de la valeur de NDVI. Ces erreurs peuvent être le fait d'une erreur de transmission produisant des valeurs nulles sur les deux canaux ou encore un effet directionnel.

paramétriques sont très sensibles aux valeurs aberrantes. Un filtrage du signal temporel, par exemple en considérant des STISc est donc souvent préalablement effectué.

D'autres auteurs proposent des méthodes de filtrage local, éventuellement accompagnées d'un filtrage autorégressif. On peut ainsi appliquer un filtre autorégressif à moyenne mobile (ARMA). [Van Dijk et al., 1987, Hird and McDermid, 2009] notent cependant un inconvénient des méthodes à moyenne pondérées qui ont tendance à altérer la forme et l'amplitude de la courbe. Par ailleurs, ces méthodes sont sensibles au bruit négatif lié aux effets atmosphériques et à la neige. Afin de réduire ces effets, [Filipova-Racheva and Hall-Beyer, 2000] proposent de combiner un filtre autorégressif à médiane mobile avec un filtre ARMA. [Ma and Veroustraete, 2006] proposent une méthode d'itération de la valeur moyenne. Cette méthode présente aussi l'inconvénient de mal restituer les transitions rapides en début de période montante, et en fin de sénescence [Hird and McDermid, 2009].

[Chen et al., 2004] propose un filtrage de type Savitzky-Golay. Ce type de filtrage calcule une valeur lissée en chaque point à l'aide d'une régression polynômiale locale. Cette méthode de lissage permet, contrairement aux méthodes de type moyenne mobile de préserver les maxima et minima locaux, ainsi que la largeur des modes.

[Hird and McDermid, 2009] proposent une comparaison empirique de six méthodes : deux d'entre elles sont des méthodes d'ajustement de courbes : ajustement par fonction logistique double (DL) [Jonsson and Eklundh, 2002], ajustement par gaussienne asymétrique [Beck et al., 2006], et les autres sont le filtre modifié de Savitzky-Golay [Chen et al., 2004], le filtre itératif de la valeur moyenne (MVI) [Ma and Veroustraete, 2006], un filtrage ARMD(3) suivi d'un filtrage ARMA(5) [Filipova-Racheva and Hall-Beyer, 2000], et le filtre *4253H, twice* [Velleman, 1980]. [Hird and McDermid, 2009] conclut sur la supériorité des deux méthodes de filtrage par ajustement de courbes : gaussiennes asymétriques, et fonction logistique double.

Les approches de filtrage, locales ou par ajustement de courbes supposent de façon plus ou moins explicite un modèle pour la phénologie. En conséquence, les méthodes dépendent fortement du type de données considéré. Ainsi, des forêts ou prairies dans des zones continentales à moyennes latitudes sont caractérisées par un seul mode par année, alors que les biomes des zones subtropicales en présentent généralement plusieurs, que les biomes tropicaux ou désertiques présentent peu de saisonnalité, et enfin, les biomes agricoles qui peuvent être très diverses en fonction de la rotation des plantes, des cultures d'hiver et d'été, etc. De même, ces méthodes dépendent fortement du type d'indice de végétation considéré. De façon générale, ces méthodes de filtrages sont sensibles aux valeurs aberrantes, et les méthodes locales nécessitent une densité importante des données pour ne pas trop lisser.

Classification spatiale à partir de STIS-CR L'information temporelle est utile pour classer les types de végétation de façon précise. En effet, différentes plantes sont caractérisées par différentes signatures temporelles liées à différents paramètres de végétation comme les dates de début et de fin de périodes végétatives, la perte des feuilles, la foliation, la défoliation, etc. Ainsi, dans [Bartalev et al., 2000], une image de paramètres phénologiques est extraite de la STISc sur laquelle une classification non supervisée par l'algorithme ERDAS-ISODATA est ensuite effectuée. Les paramètres phénologiques sont les valeurs médianes entre deux périodes supposées correspondre au printemps et à l'hiver, ainsi que la moyenne sur une période supposée correspondre à la maturité. De forts *a priori* sont ici supposés sur les données. Cette idée d'exploiter des paramètres phénologiques pour distinguer les différentes cultures est aussi exploitée dans [Bouzidi et al., 2003]. L'objectif est alors de déterminer s'il y a eu changement d'utilisation du sol d'une année sur l'autre

sur une zone agricole. Les données spatio-temporelles sont des STIS-CR NOAA/HVHRR, et un modèle de mélange linéaire des contributions d'un pixel est alors supposé (voir ci-dessous, eq A.2). Les contributions sont supposées invariantes d'une année sur l'autre en supposant la forme des parcelles constantes. Six paramètres phénologiques sont extraits des courbes, puis une analyse en composante principale est effectuée de façon à réduire l'espace et obtenir des caractéristiques décorréelées et donc plus discriminantes. Un apprentissage de ces paramètres est alors effectué sur une partie des courbes de la première année puis généralisé à toutes les courbes de la première année et de la deuxième année.

[Aurdal et al., 2005] propose d'intégrer les *a priori* phénologiques en utilisant les chaînes de Markov cachées afin d'obtenir une classification. Les évolutions des pixels sont ainsi modélisées par une succession d'états phénologiques, et les pixels sont classifiés suivant la probabilité de produire une certaine série d'états.

Le modèle d'analyse de mélange spectral linéaire (SMA) [Horwitz et al., 1971, Settle and Drake, 1993, Adams et al., 1995] a été étendu dans [Piwowar et al., 2006] à une méthode d'analyse de mélanges temporels linéaires (TMA) de façon à obtenir des images partielles de contributions caractérisées par des comportements temporels particuliers, et interprétées par des considérations phénologiques. Dans ce cas, cependant, la localisation des différentes contributions demeure inconnue.

Nous avons vu au §A.2.4.1 différentes méthodes de débruitage des signaux temporels d'indice de végétation. Cette approche consiste à considérer comme un bruit les autres contributions que la végétation du pixel moyen de 1km^2 dont on dispose. Une autre approche consiste à estimer ces différentes contributions. C'est l'approche généralement adoptée dans un but de classification de la couverture du sol.

Ainsi, [Bouzidi et al., 1998] propose une approche de fusion d'information dynamique à basse résolution avec de l'information structurelle statique. La façon la plus simple consisterait à appliquer des méthodes de fusion multi-résolution image par image de façon à obtenir une STIS à meilleure résolution. Cependant, cette approche est très coûteuse en temps de calcul et ne permet pas d'obtenir une seule classification de la zone. Une intégration de signaux temporels HR reconstruits à partir des signaux CR est alors proposée. Un modèle de mélange linéaire des contributions des pixels des parcelles à haute résolution spatiale est considéré. En notant $r_i(t)$ la réflectance au temps t du pixel i de l'image à basse résolution, et $R_j(t)$ la réflectance de la parcelle j de l'image à haute résolution au temps t , et p_{ij} la proportion de pixels HR de la parcelle j présents dans le pixel CR i , on a alors :

$$r_i(t) = \sum_{j=1}^N p_{ij} R_j(t) \quad (\text{A.2})$$

Un apprentissage des paramètres p_{ij} est effectué sur une partie d'une image HR à partir d'information sur l'occupation du sol et un ajustement par moindres carrés permet ensuite d'obtenir les réflectances HR pour tous les temps. Enfin, un apprentissage sur les classes des parcelles HR à partir des signaux temporels $R_j(t)$ précédemment estimés est effectué par un processus de minimisation déterministe de descente de gradient. La distance minimisée est une somme des distances des réflectances en chaque temps. La classification n'est donc pas effectuée dans l'espace multi-temporel, mais conjointement dans les différents espaces temporellement localisé (TL).

Le fait que les contributions des parcelles aux pixels CR soient apprises sur une seule image puis imposées aux autres est gênant car toutes les parcelles ne sont pas distinguables aux mêmes instants. [Robin et al., 2008] propose donc d'estimer les contributions des par-

celles en utilisant conjointement toutes les images de la séquence. L'information structurelle HR utilisée dans ce cas est une segmentation multi-échelle d'une image SPOT à un certain instant. Elle utilise pour cela une approche bayésienne par maximum a posteriori optimisé par un algorithme de recuit simulé. Notons qu'une hypothèse d'indépendance temporelle des différentes observations par rapport à la classification recherchée est effectuée de façon à éviter une étape d'apprentissage ou l'introduction d'*a priori*. Ici encore, il s'agit donc d'une classification jointe des pixels dans les espaces TL.

Cette méthode a été étendue dans [Robin et al., 2009] pour effectuer de la détection de changement. Une méthode de détection de changement HR à partir d'une classification HR et d'une STIS-CR basée sur une approche *a contrario* est ainsi proposée. Le principe est de vérifier la cohérence de la classification pour les différences images de la STIS-CR et de détecter les incohérences en fixant un seuil de fausses alarmes. L'indicateur de changement considéré est l'erreur d'estimation d'un pixel-CR mélangé à partir de la classification HR. Le nombre de fausses alarmes est proportionnel à la probabilité qu'un modèle naïf (gaussien uniforme sur l'image) fournisse une erreur d'estimation des pixel-CR inférieure au modèle structuré. Cette approche, très prometteuse, est donc complètement automatique. Notons cependant que les changements sont détectés par rapport à une partition de référence, et ne permettent donc pas de description évolutive des changements structurels comme une *vraie* STIS-HR le permet. Par ailleurs, les changements sont détectés sur des domaines correspondants aux pixels-CR, et il en résulte une précision géométrique réduite.

[Mercier et al., 2005] propose une analyse des changements de couverture des sols cultivables à haute résolution à partir d'une séquence d'images de SPOT VEGETATION d'une durée de trois ans à fréquence journalière. L'information à haute résolution spatiale provient d'images SPOT HRVIR acquises aussi pendant cette période mais pour une fréquence temporelle de 3 mois. De plus, de l'information temporelle exogène sous forme de carte de changements potentiels à haute résolution est incorporée. L'estimation des coefficients de mélange des pixels HR, ainsi que de leur localisation à haute résolution est effectuée grâce à des cartes de changements à organisation propre (SOM) de Kohonen. Ce type de réseaux de neurones permet non seulement d'apprendre des modèles, mais en plus de les organiser par similarité dans une grille bidimensionnelle. Une distance particulière entre la couche de Kohonen et la couche d'entrée est considérée de façon à introduire de l'information sous différentes formes : images HR et signatures spectrales CR. Cette méthode permet donc de combiner de l'information HR et de l'information multitemporelle CR dans des cas où les rapports de résolution peuvent être importants.

Enfin, une autre approche a récemment été proposée pour analyser les évolutions des pixels dans les STIS. Le principe est de discrétiser l'espace des valeurs de façon à obtenir un ensemble de séquences de symboles par pixels. La méthode a été introduite pour une application sur des images à basse résolution de météosat [Julea et al., 2006], puis a été appliquée sur la séquence ADAM [Julea et al., 2008]. Ces méthodes sont très prometteuses car elles permettent une analyse des évolutions sans *a priori* et permettent donc d'analyser des scènes hétérogènes. Cependant, l'analyse au niveau pixellaire ne permet pas de mettre en évidence des objets. Aussi, nous avons proposé dans [Le Men et al., 2008] une méthode permettant d'exploiter à la fois les structures spatiales des STIS extraites par segmentation spatiale, et l'information temporelle des évolutions des pixels. Plutôt qu'une fusion des évolutions caractéristiques des pixels et des partitions spatiales, nous proposons dans cette thèse d'appliquer ces méthodes de recherche de motifs directement aux séquences d'objets spatiaux caractérisant des objets spatio-temporels.

Nous avons présenté dans cette section différentes méthodes exploitant l'information

temporelle pour classifier les types d'occupation du sol. Certaines exploitent un *a priori* phénologique sous forme de modèle probabiliste ou en utilisant des paramètres caractéristiques. Ces méthodes sont cependant perturbées par la faible résolution spatiale qui induit un bruit dans les signaux temporels. Des approches visant à reconstruire une information à haute résolution spatiale et temporelle ont donc été proposées. Ces approches montrent bien l'intérêt des STIS-HR pour lesquelles nous proposons une méthode d'analyse dans cette thèse. Notons cependant une différence importante par rapport aux hypothèses sous-jacentes à ces méthodes : l'information structurelle de haute résolution consiste généralement en une segmentation d'une image à haute résolution. La structure est donc supposée invariante dans le temps. Dans une vraie STIS-HR, l'information structurelle varie. Choisir une seule partition à un instant afin de décrire l'évolution temporelle des régions est donc réducteur. En effet, pour des cultures suffisamment variées, les phases de maturation, spécifiques à une culture, et de dormance, commune à de nombreuses cultures, peuvent être décalées temporellement de sorte qu'il est nécessaire de considérer plusieurs temps pour distinguer correctement les parcelles. Il est donc nécessaire de prendre en compte ces changements structurels pour une description fidèle à la réalité des STIS-HR.

A.2.4.2 Extraction de structures spatio-temporelles (ST) sur des critères informationnels sur des STIS-HR

[Gueguen and Datcu, 2007] proposent une méthode permettant d'extraire l'information pertinente de la STIS à partir du principe du goulot d'étranglement de l'information (Information Bottleneck). Cette méthode permet d'estimer conjointement le modèle et ses paramètres lors d'une classification, et ainsi, d'éviter la propagation d'erreurs liées aux approches bayésiennes à deux niveaux comme celles utilisées dans [Datcu et al., 1998, Schroder et al., 1998]. Les modèles considérés sont des champs de Markov tridimensionnels (la 3^{ème} dimension étant le temps) de type Gauss-Markov (GMRF) et autobinominal. Une extension de cette méthode permettant de prendre en compte des régions a été proposée dans [Gueguen, Le Men and Datcu, 2006]. Dans cet article, deux approches ont été comparées. L'une suppose une invariance structurelle et considère une seule segmentation pour toute la séquence. En chacune de ces régions, et pour tous les temps, les paramètres GMRF sont calculés, et l'ensemble de ces séquences est classifié en utilisant le principe IB. La segmentation commune est obtenue à partir d'une segmentation LPE vectorielle sur l'ensemble de la STIS. La finesse de cette partition est réglée automatiquement en optimisant les niveaux d'immersion de cette segmentation-compromis ainsi que ceux des partitions LPE de chacune des images par un processus itératif minimisant la variation d'information entre la partition-compromis et chacune des partitions de la séquence [Giros, 2006]. L'autre méthode suppose en revanche une invariance structurelle de la séquence, et les différentes segmentations obtenues par l'algorithme qu'on vient de décrire sont utilisées. Un graphe d'adjacence temporelle reliant les régions successives sur des critères de minimisation de la variation de l'information est construit, et les séquences de textures temporellement localisées sont classifiées par le même principe du goulot d'étranglement de l'information. Cette méthode permet de mettre en évidence des structures particulières. La méthode de segmentation utilisée permet d'adapter la finesse de chaque image. Cependant, le paramètre de finesse considéré est spatialement global. Cette méthode n'est donc pas adaptée à des évolutions locales de radiométries faisant évoluer les contrastes localement. Nous proposerons dans ce manuscrit une méthode de segmentation permettant de prendre en compte la localité spatiale des changements de contrastes.

Dans [Gueguen and Datcu, 2008], la STIS est considérée comme une base de données et une méthode de compression sans pertes de cette base est proposée. Pour ce faire, chaque structure ST est décrite dans un message en deux parties, l’une décrivant le modèle, et l’autre la partie aléatoire. Les modèles sont ensuite regroupés de façon à être mis en facteur pour les structures les partageant. C’est cette mise en facteur des modèles qui permet la compression. Par ailleurs, ces modèles sont utilisés comme indices sur la base de données à partir desquels on peut effectuer des requêtes. On obtient de cette façon une compression permettant de préserver des indices pertinents sur les données compressées.

A.2.4.3 Apprentissage supervisé de structures spatio-temporelles sur des STIS-HR

La scène est constituée d’objets, et les principaux changements caractérisant ces objets sont des changements radiométriques. [Heas and Datcu, 2005] adoptent une approche pixellaire et proposent un modèle pour ces changements de façon à mettre en évidence des structures spatio-temporelles intéressantes du point de vue de l’utilisateur. Cette approche a deux étapes : la première, dans laquelle des évolutions significatives sont modélisées sous la forme de suites de classes TL, et la seconde, dans laquelle on fait intervenir un utilisateur afin d’ajouter de la sémantique. Les évolutions significatives sont obtenues à l’aide d’une quantification vectorielle d’une version réduite de l’espace multidimensionnel. La première étape consiste donc à effectuer une réduction de l’espace multidimensionnel qui est très creux (6 millions de points dans un espace à 38 dimensions pour la STIS ADAM). Cette réduction consiste en une analyse en composantes principales suivie d’une poursuite de projection. A ce stade, chacun des six millions de pixels correspond donc à une évolution dans un espace de dimension réduite. Une quantification vectorielle de cet espace est ensuite effectuée via un modèle de mélange de k gaussiennes où k est estimé à l’aide d’un critère de minimisation de la longueur de description (MDL). Afin de permettre une représentation intelligible de ces évolutions à présenter à un utilisateur, il faut ensuite repasser dans l’espace initial. De plus, afin de permettre à l’utilisateur de s’intéresser à des comportements de longueur inférieure à la durée de la STIS, une simple projection dans l’espace temporellement localisé (TL) n’est pas suffisante. Une quantification vectorielle est donc effectuée dans chaque image et les évolutions de l’espace multidimensionnel sont approchées par des graphes reliant ces valeurs discrétisées. Cette approximation est effectuée en seuillant, pour chaque instant, la probabilité conditionnelle de chaque valeur discrétisée sachant la projetée de l’évolution multi-temporelle. Un modèle interactif de type KIM (Knowledge driven Image information Mining) [Datcu et al., 2003] pour apprendre la sémantique des structures spatiotemporelles de la STIS intervient ensuite [Heas and Datcu, 2005, Héas, 2005].

Cette méthode permet d’apprendre des structures temporelles de durée variable, pertinentes du point de vue de l’utilisateur. Les résultats, présentés dans [Héas, 2005] montrent la pertinence de cette approche. Par ailleurs, on remarque que les structures mises en évidence correspondent généralement à des ensembles de régions spatiales (bruitées d’un bruit poivre-et-sel). La redondance spatiale présente dans les données ressort donc effectivement dans le résultat, mais n’a pas été exploitée explicitement par le modèle. Enfin, cette approche fait intervenir un utilisateur et présente donc les avantages et les inconvénients classiques liées aux approches semi-supervisées. A savoir que les résultats correspondent généralement mieux à l’utilisation particulière souhaitée par l’utilisateur, et par conséquent, modulo un apprentissage particulier à chaque application, l’approche peut être

utilisée pour des applications diverses. En revanche, cet apprentissage nécessaire peut être long, particulièrement dans le cas des séquences qui sont complexes. Une approche complètement automatique, d'un autre côté, permettrait de mettre en évidence des structures particulières d'intérêt que l'utilisateur n'avait peut-être pas remarqué et qui n'avaient par conséquent pas été apprises par le système.

A.2.4.4 Détection de changements

La détection de changements est le processus d'identification d'états distincts d'une zone en l'observant à des dates différentes. On distingue deux types de détections de changements. Celles qui fournissent une information binaire de type *changement* ou *non changement*, et celles qui distinguent différents types de changements.

Détection binaire de changements La détection binaire de changements est composée de deux étapes : la construction d'un indicateur de changement, et une prise de décision sur l'existence ou non de changements.

Obtention d'indices de changements Il existe tout d'abord des caractéristiques de changements ponctuels. Ainsi, [Muchoney and Haack, 1994] considèrent une différence d'images. [Prakash and Gupta, 1998] considère le rapport d'images. Ces méthodes sont néanmoins sensibles à l'intercalibration radiométrique entre les images, liée aux changements atmosphériques, ainsi qu'aux éventuels changements de capteur. [Singh, 1986, JHA and Unni, 1994] utilisent donc plutôt la différence avec une *compensée en radiométrie*. Enfin, il existe des méthodes d'analyse vectorielle de changements (CVA). Cette méthode caractérise les changements à partir de l'amplitude et la direction des vecteurs de changements spectraux [Lambin and Ehrlich, 1996]. Ces méthodes ont trouvé un regain d'intérêt avec la très haute résolution car Bovolo a montré qu'on pouvait détecter et filtrer les bruits de recalage géométrique dans cet espace [Bovolo and Bruzzone, 2007, Bovolo et al., 2008]. Enfin, [Rignot and van Zyl, 1993] proposent de caractériser les changements sur un voisinage par la moyenne locale.

Une autre catégorie d'approches consiste à se placer dans un espace transformé. Ces méthodes regroupent par exemple l'analyse en composantes principales. Cette approche permet de réduire la redondance des images et donc de focaliser sur les différences. L'espace des caractéristiques considéré peut être issu de deux dates et un canal [Wang, 1993], plusieurs dates, un canal, ou plusieurs dates et plusieurs canaux [Fung, 1990]. Cependant, les différentes dimensions sont difficiles d'interprétation, aussi, [Collins and Woodcock, 1996] proposent une transformation dans un espace *brillance-verdure-humidité* de façon à interpréter plus facilement les changements ayant lieu dans chacune des bandes. [Collins and Woodcock, 1994] propose une transformation par orthogonalisation de l'espace par la méthode de Gram-Schmidt initialisée par un choix particulier de pixels. [Ridd and LIU, 1998] proposent la *méthode du χ^2* dont l'image de changement est l'image de différence centrée et normalisée par la variance. Enfin, la méthode de séparation spectrale permet aussi de produire un ensemble d'images, *images fractions*, qui peuvent être comparées individuellement aux différentes dates [Adams et al., 1995].

Les méthodes citées jusqu'à présent sont particulièrement adaptées à une détection de changements dans des images à résolution grossière. Aux résolutions maintenant existantes, d'autres méthodes doivent être considérées. En effet, plus la résolution augmente, plus la variabilité spectrale interne augmente, et plus la variabilité spectrale externe diminue. Deux

approches permettent de s'adapter à cette particularité des images à haute résolution. L'une propose d'exploiter la variabilité intra-classe qui devient une nouvelle caractéristique de la classe. Ce sont les méthodes utilisant des textures [De Martino et al., 2003]. Et l'autre consiste à réduire la variabilité intra-classe en considérant des valeurs moyennes sur des objets plus grands.

Les méthodes de textures sont particulièrement adaptées aux images SAR qui présentent un bruit multiplicatif. [Inglada et al., 2007] propose d'aller plus loin que [Rignot and van Zyl, 1993] dans l'estimation des différences des lois de génération des pixels de chaque image en utilisant des statistiques d'ordre supérieur calculées par des développements de cumulants. Ces méthodes locales dépendent du choix d'une fenêtre d'analyse, et l'auteur propose donc un indicateur de changements à plusieurs échelles, le *profil de changement multi-échelle*. [Mercier et al., 2007] proposent de modéliser les changements de modalité de façon explicite de façon à n'être sensible qu'aux *autres* changements. L'auteur utilise pour cela la théorie des copules qui lui permettent de simuler la distribution de l'image qu'aurait produit le deuxième mode d'acquisition sur la première scène et vice-versa.

Afin de réduire la variabilité intra-classe et ce faisant d'augmenter la variabilité inter-classe de façon à faciliter les systèmes de classification automatique, on peut, à l'image du système visuel humain, adopter une approche orientée objet. Ainsi [Bruzzone and Prieto, 2000] proposent une détection des changements au niveau des régions plutôt qu'au niveau des pixels. Dans un premier temps, les images sont segmentées indépendamment, puis les partitions obtenues sont ensuite fusionnées en calculant leur intersection. Pour chaque région de cette partition fusionnée, un vecteur de caractéristiques est considéré à chaque date. Les caractéristiques utilisées dans l'article sont la moyenne et l'entropie. La méthode de fusion des segmentations tend à sur-segmenter. [Bovolo and Bruzzone, 2005] proposent donc d'étendre cette méthode à une description multi-échelle. Lors de la fusion des segmentations, l'intersection est d'abord calculée, puis les régions sont à nouveau fusionnées selon un critère d'homogénéité de plus en plus souple qui doit être vérifié dans les deux images simultanément. La segmentation reste néanmoins très sensible aux erreurs de recalage. En effet, un simple décalage d'un pixel d'une image par rapport à l'autre aura pour effet de rajouter des régions au niveau des frontières. Dans [Bovolo and Bruzzone, 2007] la méthode de filtrage du bruit de recalage mentionnée ci-dessus est donc appliquée en pré-traitement.

[Dalla Mura et al., 2008] proposent de filtrer l'image de différence en effectuant une autoreconstruction duale de l'image, puis de restaurer l'information structurale à l'image différence en appliquant des séquences de fermetures et ouvertures par reconstruction utilisant l'image originale comme masque.

Inférence d'une classe de changements Une fois que le degré de changements a été évalué, on est amené à décider si la zone considérée a changé ou pas. Des méthodes de classification automatique ou supervisée doivent déterminer par le choix d'un seuil approprié si celui-ci a effectivement eu lieu. L'approche bayésienne est l'approche majoritairement utilisée pour fixer ce seuil. [Bruzzone and Serpico, 1997] proposent de classifier une image de différence entre deux dates par une approche basée sur la règle de Bayes. Deux classes de pixels sont considérées : les pixels où il y a des changements ω_c et ceux où il n'y en a pas ω_n . La probabilité de la valeur X de chaque pixel de l'image de différence peut être écrite comme : $p(X) = p(X|\omega_n)p(\omega_n) + p(X|\omega_c)p(\omega_c)$. Sous cette hypothèse, l'algorithme de maximisation de l'espérance, EM est utilisé afin d'estimer $p(X|\omega_n)$, $p(\omega_n)$, $p(X|\omega_c)$ et $p(\omega_c)$. Les probabilités conditionnées sont supposées suivre des lois gaussiennes,

ce qui simplifie l'algorithme à l'estimation de la moyenne et de la variance des lois.

Un autre exemple est proposé par [Smits and Annoni, 2000]. Des zones de changements sont détectées par la communication au système de la probabilité de non-détection et de la probabilité de fausse alarme. Dans le cas automatique (classification non-supervisée), la façon la plus simple de procéder est de faire une classification sur les deux dates indépendamment, puis de calculer la différence des classifications. Cette approche dépend beaucoup des performances des algorithmes de classification utilisés.

Classification de changements Une autre catégorie d'approche utilise des classifications. Ainsi, [Weismiller et al., 1977] propose d'effectuer la classification de l'ensemble des points de l'espace spectro-spatio-temporel, puis identifie les changements pixels à pixel en retournant dans l'espace image. Afin de minimiser l'effets de changements d'acquisition, [Munyati, 2000] propose une classification temporellement localisée de chacune des images suivie d'une comparaison pixel à pixel permettant de générer des matrices de transitions de classes. [Bruzzone and Serpico, 1997] proposent une classification de la probabilité jointe des valeurs de pixels consécutifs par un algorithme de maximisation de l'espérance (EM). Des classifications hybrides utilisant un masque de détection de changements permettent de focaliser l'analyse sur les zones d'intérêt et de permettre une classification plus ciblée [Macleod and Congalton, 1998]. Enfin, des réseaux de neurones peuvent être utilisés pour classifier les changements [Liu et al., 2002]. Cette approche supervisée permet d'identifier plus facilement les changements d'intérêt pour l'utilisateur.

Détection de changements pour la mise à jour de GIS Enfin, une dernière classe de méthodes s'intéresse à la détection de changement dans un but de mise à jour de GIS. La détection peut alors concerner deux images ou une image et un GIS. [Knudsen and Olsen, 2003] utilisent une classification d'une image aérienne, qu'il superpose ensuite à la base de données vectorielle de façon à obtenir les changements. Afin de faciliter la comparaison entre les classifications et de minimiser les changements géométriques, [Walter, 2004] propose de classifier les objets correspondant au GIS dans la nouvelle image. Cette méthode permet donc de détecter facilement les changements d'occupation du sol, en revanche, les changements géométriques seront difficilement détectés. Ainsi, la construction d'une maison dans une grande forêt ne sera pas détectée. [Carleer and Wolff, 2007] proposent de comparer une classification de régions obtenues par segmentation par croissance de régions à la base de données. Dans cette dernière approche, des changements structurels sont donc pris en compte, cependant, l'intérêt n'est pas sur ces changements mais sur les changements de classes en termes de proportions représentées dans une matrice de changement d'utilisation du sol.

A.2.5 Mise en correspondance d'images

Il existe différentes méthodes de mise en correspondance d'images. Nous avons vu des applications à la poursuite de cible et à la correction du mouvement dans des séquences d'images et nous nous intéressons particulièrement, ici, à la mise en correspondance pour la reconstruction de scènes tridimensionnelles. Dans ce cas, la mise en correspondance des projections des objets dans différents plans image,⁶ est contrainte par des relations

6. Le plan image est le plan perpendiculaire à l'axe optique et contenant le foyer de la caméra. Dans le cas stéréo, il existe plusieurs caméras et donc plusieurs plans image.

appelées *contraintes épipolaires* (ou quasi-épipolaire dans le cas de perspective cylindro-conique). Cette contrainte traduit le fait qu'un point dans une image appartient à une certaine droite dans l'autre image. L'espace est ainsi réduit à un espace de dimension un ce qui simplifie grandement le problème. L'estimation de la géométrie relative des deux capteurs revient à estimer six paramètres (le vecteur de déplacement entre les deux centres, et la différentes des angles de l'axe optique).

Les correspondances peuvent concerner des attributs d'objets dans les images ou des relations. Une première catégorie de méthode appelée *mise en correspondance par patrons* consiste à mettre en correspondance des blocs. Comme nous l'avons vu au §A.2.3.3, cette approche n'est pas robuste aux changements radiométriques ni aux occlusions. Nous noterons néanmoins parmi ces méthodes, les méthodes par génération et propagation d'hypothèses proposées par [Ayache and Faugeras, 1986] qui sont des méthodes qui exploitent la cohérence spatiale des appariements de façon à réduire la complexité. Une autre catégorie d'approches propose d'extraire des primitives saillantes et de les mettre en correspondance. Les primitives typiquement extraites sont des coins. De plus, ces primitives sont généralement liées de sorte que la mise en correspondance d'images se traduit par une mise en correspondance de sous-graphes. Dans ce cas, les méthodes d'estimation généralement utilisées sont la recherche d'arbres, la relaxation et la recherche de cliques maximales.

[Shapiro and Brady, 1995, Olsen, 1992, Zhang et al., 1995] proposent ainsi une estimation de cette géométrie pour la mise en correspondance d'images de stéréovision. L'estimation s'appuie sur la mise en correspondance d'un certain nombre de points caractéristiques (deux dans le cas de [Shapiro and Brady, 1995, Olsen, 1992], quatre dans [Xu et al., 1993, Nishimura et al., 1993], et huit dans [Zhang et al., 1995]). Dans cette problématique, seule des correspondances 1-1 sont considérées. Afin d'assurer cette unicité, des méthodes dites de *relaxation* sont généralement appliquées après avoir sélectionné un ensemble de correspondances candidates. Ces méthodes s'appuient généralement sur le contexte en supposant que les vraies correspondances présentent un grand nombre de vraies correspondances dans leur voisinage aussi. Une force de mise en correspondance dépendant du voisinage est ainsi considérée, et la relaxation consiste à minimiser la somme de ces forces pour toutes les correspondances. Il existe pour ce faire différentes approches. La première, *winner-take-all* [Rosenfeld et al., 1976], consiste à parcourir toutes les correspondances en ne gardant que la meilleure. Cette approche présente le principal inconvénient qu'alors, des correspondances en fin de parcours peuvent se trouver *meilleures* alors qu'elles avaient une mauvaise correspondance initialement. Une autre approche, *looser-take-nothing* consiste à parcourir itérativement l'ensemble des correspondances en supprimant pour chacune la pire correspondance [Li, 1990]. Enfin, une approche dite *some-winners-take-all* a été proposée dans [Zhang et al., 1995] dans laquelle, pour chaque correspondance, celles qui sont à la fois dans les $q\%$ meilleures correspondances et dans les $q\%$ moins ambiguës sont sélectionnées, et les autres correspondances impliquant les deux caractéristiques gagnantes sont supprimées.

Notons que les contraintes épipolaires ne s'appliquent qu'à des points de l'espace. Pour mettre en correspondance des droites comme des contours ou même des objets, ces contraintes ne peuvent donc plus être utilisées. [Horaud and Skordas, 1989] propose donc de mettre en correspondance d'abord des points en exploitant les contraintes épipolaires, puis les relations entre ces points par recherche de cliques maximales dans un graphe de correspondance. Les nœuds d'un tel graphe représentent les correspondances entre une caractéristique d'une image et une caractéristique de l'autre image, et les arcs traduisent la compatibilité de ces correspondances. Les plus grandes cliques maximales,⁷ dans le

7. Une clique d'un graphe est une composante complètement connexe : chaque nœud de la clique est lié

graphe de correspondance sont ensuite extraites. La compatibilité des nœuds du graphe de correspondance est déterminée par un ensemble de règles assurant la cohérence interne des caractéristiques considérées. Ici encore, ce point de vue diffère des mises en correspondances que l'on cherche à mettre en évidence entre les objets de la séquence ADAM car on s'intéresse à une mise en correspondance complète des objets de l'image et qu'aucune contrainte ne peut être appliquée pour déduire d'une correspondance les correspondances des autres objets.

Il existe aussi des méthodes qui consistent à extraire des cartes de distance entre les pixels correspondants dans les deux images. De telles cartes sont appelées *cartes de disparité*,⁸ [Hong and Chen, 2004, Klaus et al., 2006]. Il est à noter que des contraintes sur la position des objets sont généralement considérées de façon à réduire l'espace de recherche des correspondances d'une caractéristique.

Enfin il existe des approches utilisant des graphes modèles. Ainsi, dans [Fuchs and Le-Men, 2000], une méthode d'isomorphisme approché de sous-graphes est effectuée en s'appuyant de connaissances extérieures représentées dans un graphe modèle dont certains nœuds sont mis en correspondances avec les données. Cette approche est appliquée à la reconstruction 3D de bâtis. Nous mentionerons finalement une autre approche légèrement différente car il ne s'agit plus de mettre en correspondance des images, mais directement des graphes. Ces approches sont notamment utilisées pour mettre en correspondance un graphe modèle avec des graphes objets. Ainsi, dans [Mangin, 1995], un modèle structurel de la topographie corticale est inféré à partir d'une base de graphes relationnels attribués de cortex cérébraux. L'utilisation d'un graphe aléatoire permet de prendre en compte la variabilité des structures, et les changements topologiques sont contraints par des contraintes syntaxiques.

Dans les cas présentés ci-dessus, une mise en correspondance de sous-graphes est effectuée de façon à estimer la troisième coordonnée ou d'inférer un modèle. Dans le cas de la séquence ADAM, l'objectif est assez différent car un isomorphisme de sous-graphe ne permet de mettre en évidence que ce qui n'a pas changé or dans notre cas, on s'intéresse particulièrement à ce qui a changé et on souhaite expliquer aussi bien l'*histoire* d'un objet conservé que d'un objet changé. Ces approches ne sont donc pas directement applicables à notre cas.

A.3 Recherche de motifs fréquents

Dans cette thèse, un algorithme issu du domaine de la recherche de motifs fréquents a été utilisé. Nous présentons dans cette annexe un complément au chapitre 6. Des problématiques connexes, ainsi que les algorithmes associés y seront présentés. Néanmoins, pour des raisons qui sont présentées au paragraphe 9.1.2, ces algorithmes n'ont pas été appliqués à nos données.

à tous les autres nœuds de la clique. Et une telle clique est dite maximale si aucun nœud connecté à tous les nœuds courants ne peut être ajouté. Enfin, la taille d'une clique fait référence au nombre de nœuds qui la constituent.

8. La disparité est la différence des abscisses curvilignes le long des épipolaires. Cette grandeur est liée à la profondeur des objets.

A.3.1 Apriori

Cet algorithme, proposé par [Agrawal and Srikant, 1994], consiste en un parcours en largeur de l'arbre de préfixe permettant d'exploiter pleinement la propriété 6.2.1. Les niveaux de l'arbre sont construits complètement les uns après les autres en partant du premier niveau, ne contenant que les symboles fréquents. Pour chaque niveau, il y a une première phase de génération des candidats, et une seconde phase d'énumération de ces candidats.

La phase de génération des candidats est elle-même décomposée en deux étapes. D'abord la jonction de deux k -motifs fréquents ne différant que d'un élément. Leur jonction produit donc un $k+1$ -motif dont ils sont des sous-motifs. Ensuite, on génère les k -sous-motifs du candidats dont on vérifie l'existence dans le niveau supérieur de l'arbre afin de vérifier qu'ils sont fréquents.

Vient alors l'étape d'énumération des candidats. Dans cette étape, l'ensemble des objets de \mathcal{O} sont parcourus, et leurs k -motifs sont générés de façon à incrémenter les compteurs des k -candidats lors d'un double parcours de ces k -motifs et des k -candidats.

La connaissance des niveaux précédents complets permet de réduire le nombre de candidats à l'ensemble des motifs fréquents plus la bordure négative de l'ensemble des motifs fréquents, c'est-à-dire à l'ensemble des motifs non-frequents dont un sous-motif est fréquent.

Afin de générer les candidats et de les énumérer de façon efficace, plusieurs structures de données efficaces ont été proposées. Ainsi, [Park et al., 1995] propose de stocker l'ensemble des motifs fréquents et l'ensemble des candidats dans deux tables de hachage. L'utilisation de la table de hachage permet de plus de réduire encore le nombre de candidats issus de la première phase de leur génération (jointure). Une structure plus efficace pour le stockage des candidats a été proposée dans [Bodon, 2003] : le trie. Il s'agit d'un arbre dans lequel les symboles sont stockés sur les arcs de sorte qu'un parcours d'une branche de l'arbre depuis la racine jusqu'à sa feuille fournit un motif fréquent. Cet arbre [De La Briandais, 1959, Fredkin, 1960] a été initialement utilisé pour retrouver facilement des chaînes de caractères. De plus, dans notre cas, le trie permet d'effectuer efficacement l'étape de jointure lors de la génération de candidats car les fils candidats d'un nœud sont alors simplement obtenus par copie de ses frères situés à sa droite.

Cette approche nécessite donc autant de parcours de \mathcal{O} qu'il y a de niveaux dans l'arbre, c'est-à-dire que la taille du motif fréquent le plus long. Différentes méthodes de réduction de la base de données ont donc été proposées. Dans [Agrawal and Srikant, 1994], la base de données est transformée de façon à stocker les k -motifs fréquents contenus dans chaque objet plutôt que l'ensemble des symboles. Cette représentation présente aussi l'intérêt de faciliter la jointure, cependant, comme nous l'avons vu, la structure de trie permet d'effectuer cette étape efficacement. Une réduction supplémentaire de la base de données est proposée dans [Park et al., 1995] en enlevant dans chaque échantillon de la base de données courante, les symboles qui apparaissent dans moins de k k -candidats de l'échantillon considéré. Différentes variantes de l'algorithme Apriori ont été proposées. Parmi elles : [Bayardo Jr and Agrawal, 1999] propose d'utiliser une borne inférieure de la fréquence d'un motif calculée à partir de la fréquence de ses sous-motifs. [Brin et al., 1997] propose de diviser l'ensemble d'objets en sous-ensembles. Sur chacun de ces sous-ensembles, les 1-candidats sont comptés puis après avoir parcouru cet ensemble, des 2-candidats sont générés avant que l'ensemble des objets n'aient été parcourus, dans le sous-ensemble suivant, on compte les 1-candidats, et les 2-candidats, et ainsi de suite. [Hidber, 1999] propose une énumération *à la volée* lors du parcours de l'ensemble des objets. [Toivonen et al., 1996]

propose de réduire le nombre de parcours par échantillonnage de l'ensemble des objets. Les motifs fréquents sont énumérés sur un échantillon de l'ensemble des objets, puis la fréquence sur la base de données complète des motifs fréquents trouvés à petite échelle sont ensuite calculés. Un second parcours de l'ensemble des objets permet de vérifier qu'aucun motif fréquent n'avait été oublié dans la recherche sur l'échantillon. Enfin, [Savasere et al., 1995] propose un calcul des fréquence par intersection des listes d'identifiants d'objets couvrants associés. Stocker l'ensemble des identifiants pour chaque motif est cependant coûteux en place mémoire, et une partition de l'ensemble des objets est alors proposée. Les ensembles de motifs fréquents de chaque composante sont fusionnés dans une dernière étape.

A.4 Algorithmes en profondeur

Pour certains ensembles d'objets, le nombre de motifs fréquents est trop important, et l'arbre de préfixe ne peut alors plus être contenu en mémoire. C'est le cas particulièrement pour les ensembles de longs objets. Une alternative pour la génération des motifs fréquents est alors un parcours en profondeur de l'arbre de préfixe au cours duquel seule une partie de l'arbre doit être maintenu en mémoire.

A.5 Croissance-FP

Cet algorithme proposé dans [Han et al., 2004] propose une structure de stockage de l'ensemble des objets permettant de combiner les avantages des deux représentations verticales et horizontales. Il s'agit d'un arbre de préfixe dont les branches correspondent aux motifs des différents objets ordonnés par fréquence décroissante. Des arcs supplémentaires relient les nœuds de même symbole apparaissant dans différentes branches. Cette structure permet donc non seulement un parcours efficace des objets grâce à leur organisation dans un arbre ainsi qu'une possibilité de traiter plusieurs objets à la fois, et de plus l'information de fréquence des différents symboles est présente grâce aux arcs supplémentaires transversaux. Ces arcs sont de plus liés à une table d'en-tête contenant la fréquence de chaque symbole.

La croissance des motifs est ensuite effectuée en *projetant* récursivement les arbres-FP par rapport aux symboles les plus bas dans l'arbre. Cette projection est calculée en propageant leur fréquences le long des branches jusqu'à la racine. L'arbre est ensuite simplifié de façon à enlever les symboles non fréquents, puis cet arbre est à nouveau projeté par rapport à son dernier niveau, etc.

L'intérêt de cet algorithme par rapport à ECLAT est la représentation compressée des couvertures des motifs. Cependant, cette compression se paye par des modifications complexes de la structures aboutissant à un nombre important de déréréférences alors qu'ECLAT ne fait que des intersections.

A.5.1 Représentations condensées

Des filtrages peuvent être effectués pour Apriori et pour ECLAT permettant de ne fournir que les motifs fermés ou les motifs maximaux [Borgelt, 2003]. Des méthodes dédiées ont aussi été proposées.

A.5.1.1 Recherche de motifs fermés

Afin de réduire l'espace de stockage de l'ensemble de motifs, une solution consiste à chercher les motifs fermés. C'est-à-dire les motifs fréquents dont motif couvrant n'a la même fréquence. Le nombre de motifs à stocker est ainsi réduit, et cette approche demeure sans perte d'information dans le sens où l'on peut utiliser les motifs fermés pour déterminer de façon unique l'ensemble des sous-ensembles fréquents ainsi que leur fréquence exacte. Différentes méthodes de recherche de motifs fermés ont donc été proposées [Bastide et al., 2000, Pasquier et al., 1999, Pei et al., 2000, Wang et al., 2003a, Zaki, 2000b].

On appelle générateur du motif fermé c l'un des plus petits ensembles de symboles dont la fermeture est c , et on note \mathcal{C} l'ensemble des motifs fermés du jeu de données.

Dans [Pasquier et al., 1999], la méthode proposée consiste à trouver l'ensemble des générateurs en faisant croître leur taille niveau par niveau par un algorithme Apriori dont la deuxième étape de la génération des candidats supprime de plus tous les $(k+1)$ -ensembles dont au moins un k -ensemble n'est pas k -générateur et ne garde que les $(k+1)$ -ensembles dont aucun k -ensemble n'a la même fréquence. L'ensemble des objets est donc parcouru une fois de plus que la taille du dernier générateur. La taille du dernier générateur est inférieure au nombre de niveaux de l'arbre de préfixe pour un ensemble d'objets corrélée, et égal si l'ensemble d'objets n'est pas corrélée. Cette méthode est donc plus efficace qu'*Apriori* pour des jeux de données corrélés. Afin d'éviter de reparcourir de nombreuses fois l'ensemble d'objets, [Zaki, 2000b] propose d'utiliser le format vertical pour stocker les données. Les fréquences sont ainsi aisément calculées en intersectant les couvertures de ses sous-motifs. Afin de permettre de stocker de grosses bases de données dans ce format, un partitionnement des données en sous-problèmes indépendants est proposé dans [Zaki, 2000b]. Le partitionnement proposé ici diffère de celui de [Savasere et al., 1995] car il utilise des classes d'équivalences. Cette notion lui permet de s'assurer que les ensembles fréquents dans une partition sont fréquents aussi dans le jeu de données complet, et ainsi de s'affranchir de la deuxième phase de [Savasere et al., 1995]. La relation d'équivalence considérée est la relation d'équivalence par préfixe : $s_i \equiv_{\theta_k} s_j \Leftrightarrow s_i[1] \dots s_i[k] = s_j[1] \dots s_j[k]$. La classe d'équivalence du préfixe P est notée $[P]$. Enfin, plutôt que de procéder en deux étapes : 1) construction des générateurs, 2) calcul des fermetures, [Zaki, 2000b] propose de calculer les fermetures lors de la construction des ensembles fréquents en exploitant des propriétés vérifiées par la jonction de deux motifs.

Dans le cas où l'on dispose d'un très grand nombre d'objets, l'ensemble des motifs fermés peut être trop petit. [Kryszkiewicz, 2002, Kryszkiewicz and Gajek, 2002, Calders and Goethals, 2003, Bykowski and Rigotti, 2001, Kryszkiewicz, 2001] proposent donc d'utiliser une structure plus condensée. Le principe est d'utiliser des règles permettant de déduire la fréquence de certains motifs à partir d'autres. Ainsi, en ne stockant que ces règles, et la fréquence des motifs initiaux, les dits motifs *libres de disjonctions*, on peut déduire des fréquences de l'ensemble des motifs fréquents. [Kryszkiewicz, 2002, Kryszkiewicz and Gajek, 2002] proposent une compression encore plus importante, mais avec pertes, permettant d'obtenir des approximations de la fréquence de l'ensemble des motifs fréquents. Les règles considérées dans ce cas ne sont pas exactes, mais présentent une forte valeur de confiance (c'est-à-dire qu'elles sont presque toujours vraies) et permettent donc de déduire (une approximation de) la fréquence de certains motifs à partir d'autres.

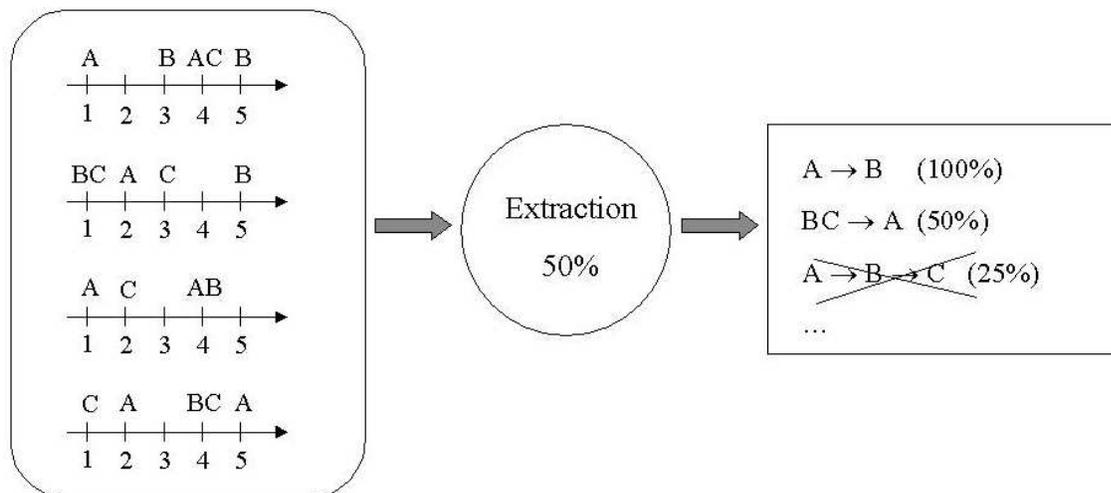


FIGURE A.1 – Illustration d’une base de données séquentielles et de l’extraction de motifs séquentiels fréquents pour un seuil de fréquence relative de 50%. [Leleu, 2004]

A.5.2 Recherche d’ensembles de symboles maximaux

Pour de très grosses bases de données, même l’ensemble des sous-ensembles fermés fréquents peut être trop important pour tenir en mémoire. On peut alors chercher les ensembles fréquents maximaux qui est un ensemble inclus dans l’ensemble des sous-ensembles fréquents fermés [Gouda and Zaki, 2001, Burdick et al., 2001, Kedem, 1998]. L’algorithme proposé dans [Gouda and Zaki, 2001] consiste en une recherche en profondeur à l’aide d’un algorithme de *retour sur trace*. De façon à élaguer efficacement l’arbre de recherche, un test vérifiant qu’aucun super-ensemble du k -ensemble courant n’a déjà été mis dans l’ensemble des sous-ensembles maximaux fréquents lors du parcours antérieur d’une autre branche. [Burdick et al., 2001] utilise aussi une recherche en profondeur, mais utilise un format vertical conjointement avec une technique de compression de façon à accélérer le calcul de la fréquence. [Kedem, 1998] recherche des ensembles fréquents maximaux avec une approche mixte montante et descendante. L’approche descendante utilise la réciproque de la propriété d’anti-monotonie : si un ensemble est fréquent, alors tous ses sous-ensembles sont fréquents, et il est inutile de les traiter dans l’approche montante.

[Agarwal et al., 2000] propose un algorithme appelé *DepthProject* utilisant un parcours en profondeur de l’arbre de recherche particulièrement adapté aux jeux de données constitués de peu d’échantillons, mais dont les motifs maximaux sont longs.

A.5.3 Recherche de motifs séquentiels fréquents

Plusieurs applications supposent un ordonnancement supplémentaire des symboles. C’est le cas par exemple de données temporelles financière par exemple, de données textuelles, des séquences de DNA, etc. Les objets considérées dans ce cas sont des séquences. La base de données peut être constituée d’une seule séquence d’événements dans laquelle on recherche des épisodes fréquents [Mannila et al., 1995], ou plus généralement de plusieurs séquences d’événements. Plusieurs événements peuvent survenir à une même date. Ainsi, dans la figure A.1, dans la première séquence, les événements A et C surviennent simultanément à l’instant 4. Une sous-séquence $\alpha_1 \rightarrow \dots \rightarrow \alpha_n$ est une sous-séquence de $\alpha'_1 \rightarrow \dots \rightarrow \alpha'_m$

ssi il existe une suite d'indices $1 \leq i_1 \leq \dots \leq i_n \leq m$ telle que $\alpha_1 \subseteq \alpha'_{i_1}, \dots, \alpha_n \subseteq \alpha'_{i_n}$. Un motif séquentiel p est dit fréquent s'il est une sous-séquence d'au moins σ_{\min} séquences, où σ_{\min} est, comme précédemment, un paramètre fourni à l'algorithme par l'utilisateur. Dans le cas de la recherche d'épisodes dans une unique longue séquence, un motif séquentiel est dit fréquent si il est une sous-séquence au moins σ_{\min} fois. La figure A.1 illustre le principe de la recherche de motifs séquentiels.

L'espace de recherche est dans ce cas beaucoup plus important pour les motifs séquentiels que pour les motifs désordonnés sans répétition car il faut alors considérer l'ensemble des permutations possibles des symboles, et non plus "simplement" les combinaisons. Ce problème a été introduit dans [Agrawal and Srikant, 1995]. Comme pour les recherches de motifs, les séquences fréquentes peuvent être organisées dans un arbre de préfixe, et afin de générer les séquences fréquentes, cet arbre de préfixe peut être parcouru en largeur ou en profondeur. La base de données peut aussi être représentée de façon horizontale [Agrawal and Srikant, 1995, Masegla et al., 1998, Srikant and Agrawal, 1996] ou verticale [Ayres et al., 2002, Mannila et al., 1995, 1997, Zaki, 1998, 2001, Leleu et al., 2003]. Ces algorithmes consistent, comme pour le cas de motifs désordonnés, en deux étapes, une étape de génération de motifs séquentiels candidats, et une étape d'énumération de ces candidats. Comme précédemment, la génération de candidats consiste en une jonction de deux motifs de même $k-1$ -préfixe. L'utilisation d'un format vertical permettant de stocker les listes d'occurrences des différents motifs permet d'obtenir les listes d'occurrences des jointures par intersection des listes d'occurrences des k -motifs joints. Notons que dans le cas des séquences, une occurrence est caractérisée par deux identifiants : celui de la séquence qui le contient, et l'indice temporel dans cette séquence. [Leleu et al., 2003] propose une extension de ces modèles permettant de gérer de façon plus efficace les séquences contenant plusieurs répétitions. En effet, dans le format vertical classique, les tables associées aux symboles peuvent être très redondantes. L'auteur propose donc une version comprimée de ces séquences permettant non seulement un gain d'espace de stockage, mais surtout des opérations d'intersections beaucoup plus rapides puisqu'elles concernent de plus petits vecteurs.

Comme nous l'avons noté, l'espace de recherche pour les séquences est considérablement plus important que pour les motifs désordonnés. Afin de faire face à cette limitation, des auteurs ont proposé d'introduire des contraintes imposées par l'utilisateur et utilisées lors de la génération des motifs séquentiels fréquents pour réduire l'espace de recherche [Srikant and Agrawal, 1996, Masegla et al., 1998, Garofalakis et al., 1999, Zaki, 2000a, Albert-Lorincz, 2003, Leleu et al., 2003]. De telles contraintes sont par exemple l'intervalle temporel maximal ou minimal entre deux événements et la fenêtre temporelle maximale (intervalle entre l'instant du premier événement et du dernier événement d'un motif séquentiel).

Des approches par projections ont aussi été proposées [Pei et al., 2001]. Enfin, dernièrement, [Gouda et al., 2007] propose d'encoder les séquences grâce à une décomposition en facteurs premiers.

A.6 Recherche dans des graphes

Le domaine de la recherche exacte de structures fréquentes s'est dernièrement étendu à la structure des graphes avec des applications principalement pour le web et la bioinformatique.

A.6.1 Recherche de traversées fréquentes dans des graphes

La recherche de séquences fréquentes a aussi été utilisée pour analyser des bases de données de séquences de clics [Chen et al., 1998, Nanopoulos and Manolopoulos, 2000, Gudes et al., 2006, Demiriz, 2004]. Afin de prendre en compte la structure du web, [Nanopoulos and Manolopoulos, 2000, Gudes et al., 2006, Demiriz, 2004] considèrent un graphe dirigé dont les nœuds représentent les pages visitées, les arcs les hyperliens entre ces pages, et une traversée dans ce graphe, les navigations de l'utilisateur sur le site. En effet, les algorithmes précédemment proposés [Chen et al., 1998] permettent de retrouver des sous-séquences de nœuds consécutifs. On appelle de telles séquences des *sections*. [Nanopoulos and Manolopoulos, 2000] propose de relaxer cette contrainte de consécuitivité qui n'est pas toujours pertinente et les méthodes qui l'utilisent souffrent donc de sensibilité au bruit. En effet, un internaute peut taper une url dans le navigateur pour aller voir une certaine page alors qu'il est en train d'en regarder une autre sans que ces deux pages soient liées. [Nanopoulos and Manolopoulos, 2000] propose en revanche d'utiliser un graphe dirigé représentant les liens entre pages de façon à contraindre les séquences pertinentes, et cherche des *sous-chemins* fréquents. Dans ce cas, à moins que les deux pages visitées par l'internaute ne soient effectivement liées, on ne les considèrera pas dans une même séquence.

Définition A.6.1.1 Sous-chemin: *Etant donné un chemin $p = p[1] \dots p[p]$, dans un graphe \mathcal{G} , on appelle sous-chemin de p la séquence $m = m[1] \dots m[m]$ pour laquelle :*

- $\forall m[i] \in m \Rightarrow m[i] \in p$
- $m[i] = m[k]$, et $m[j] = p[l]$ et $i < j \Rightarrow k < l$, $\forall i < j$

L'algorithme proposé dans [Nanopoulos and Manolopoulos, 2000] pour trouver les séquences fréquentes contraintes par le graphe est un algorithme de type Apriori. Cependant, la génération de candidats diffère car les candidats qui nous intéressent ici ne sont plus seulement des ensembles de symboles candidats à être fréquents, mais des sous-chemins candidats à être fréquents, c'est-à-dire qu'ils doivent en plus correspondre à un sous-chemin dans le graphe, et que chacun de leurs sous-chemins doivent être fréquents. Pour ce faire, des $(k+1)$ -candidats sont générés par jonction d'un k -chemin fréquent l_i dont le dernier nœud est $l_i[k]$, et d'un k -chemin fréquent dont les $(k-1)$ premiers nœuds sont les $(k-1)$ derniers nœuds de l_i , et dont le dernier nœud est un voisin de $l_i[k]$ dans le graphe représentant la structure du graphe. Finalement, le calcul des fréquences des candidats diffère aussi car pour générer les sous-chemins existants d'un objet, on commence par générer ses C_k^n sous-ensembles puis pour chacun d'eux, on vérifie qu'il s'agit bien d'un sous-chemin existant, et si c'est le cas, on incrémente sa fréquence. La recherche de traversées fréquentes a été récemment étendue dans [Lee and Park, 2007] à des graphes pondérés. De tels poids peuvent représenter l'importance d'une page web. En conséquence, la propriété de fermeture vers le bas est perdue, et l'algorithme utilise alors une borne de la fréquence des sous-chemins.

A.6.1.1 Recherche de sous-graphes fréquents

Le but est ici de trouver les sous-graphes fréquents dans une base de graphes. Trois types d'approches existent pour ce type de problèmes : une approche consistant à étendre les méthodes de recherche de type Apriori à ce problème en ajoutant successivement des sommets à des sous-graphes fréquents [Inokuchi et al., 2000], d'autres méthodes, aussi basées sur une heuristique *a priori*, mais qui agrandissent les sous-graphes fréquents en rajoutant des arcs [Kuramochi and Karypis, 2001, 2004], et des approches gloutonnes qui

ne garantissent pas d'obtenir tous les sous-graphes, mais sont relativement rapides, et permettent une mise en correspondance approchée [Holder et al., 1994, Gonzalez et al., 2001, Jonyer et al., 2002].

Dans [Kuramochi and Karypis, 2001, 2004], un k -graphe est un graphe contenant k arcs. Une heuristique de type Ariori est utilisée pour réduire la recherche, en construisant les $(k+1)$ -sous-graphes fréquents à partir de paires de k -sous-graphes fréquents qui ont un $(k-1)$ -graphe en commun. Ce dernier est appelé le noyau. Elles utilisent donc la propriété de fermeture vers le bas du treillis de sous-graphes fréquents : les $(k-1)$ -sous-graphes d'un k -sous-graphe fréquent sont fréquents. L'algorithme se décompose en une étape de génération de k -sous-graphes candidats à partir de l'ensemble des $(k-1)$ -sous-graphes fréquents. Puis pour chacun des k -sous-graphes candidats, et pour chaque graphe-échantillon de la base, on regarde si le k -sous-graphe est un sous-graphe du graphe-échantillon. Si c'est le cas, on incrémente la fréquence associée au k -sous-graphe. On calcule ainsi la fréquence de chaque k -sous-graphe candidat, et on peut en déduire l'ensemble des k -sous-graphes fréquents.

Des comparaisons de graphes sont donc nécessaires à plusieurs reprises au cours de l'algorithme. Par exemple, lors de la vérification de la condition de fermeture vers le bas : lorsqu'on vérifie que chaque $(k-1)$ -sous-graphe d'un k -graphe est fréquent (appartient à F^k) avant de l'ajouter à C^k , et lors de la génération de l'ensemble des candidats pour vérifier que le k -graphe courant n'est pas déjà dans C^k . Cependant, l'isomorphisme de graphes est un problème qui nécessite la comparaison des différentes représentations d'un graphe obtenues par permutation des sommets et des arcs. [Kuramochi and Karypis, 2004] propose d'éviter ce calcul en considérant une labellisation canonique des graphes consistant à classer les nœuds par degré croissant, et à degré égal, par label croissant. Une unique chaîne de symboles est ainsi associée à chaque graphe. Une fois que le label canonique d'un graphe a été calculé, tester l'égalité revient finalement à comparer deux chaînes de caractères, et peut donc être effectuée à de nombreuses reprises à coût réduit.

A.7 Généralités sur les graphes

Définition A.7.0.2 Partition: Une partition d'un espace D est une collection de sous-ensemble disjoints et non vide $\{R_i\}_{i=1}^{|\mathcal{P}|}$ de D dont la réunion est D .

Définition A.7.0.3 Partition en composantes connexes : Une partition en composantes connexes est une partition dont les composantes sont connexes, c'est-à-dire qu'il existe un chemin reliant n'importe quelle paire de points de la composante ne passant que par des points de la composante. Par la suite nous ne nous intéresserons qu'aux partitions connexes même s'il nous arrivera d'omettre de le préciser.

Définition A.7.0.4 Graphe: Un graphe est composé de deux ensembles, l'ensemble des arêtes (\mathcal{H}) et l'ensemble des sommets (\mathcal{U}). L'ensemble des sommets est simplement une collection d'étiquettes qui permettent de distinguer un sommet d'un autre. L'ensemble des arêtes est constitué de paires non ordonnées d'étiquettes de sommets.

Définition A.7.0.5 Sous-Graphe: Un sous-graphe $G' = (\mathcal{U}', \mathcal{H}')$ du graphe $G = (\mathcal{U}, \mathcal{H})$ est un graphe dont les sommets sont des sommets de G et dont les arcs sont des arcs de G : $\mathcal{U}' \subset \mathcal{U}$, et $\mathcal{H}' \subset \mathcal{H}$.

Définition A.7.0.6 Sous-graphe induit: Pour un sous-ensemble de sommets $\mathcal{U}' \subset \mathcal{U}$, le sous-graphe de G induit par \mathcal{U}' est le graphe $G' = (\mathcal{U}', \mathcal{H}(\mathcal{U}'))$ dont l'ensemble des sommets est \mathcal{U}' et l'ensemble des arêtes $\mathcal{H}(\mathcal{U}')$ est formé de toutes les arêtes de G ayant leurs deux extrémités dans \mathcal{U}' . Autrement dit, on obtient G' en enlevant un ou plusieurs sommets au graphe G , ainsi que toutes les arêtes incidentes à ces sommets.

Définition A.7.0.7 Graphe connexe: Un graphe connexe est un graphe dans lequel chaque paire de sommets est reliée par une chaîne. Un graphe qui n'est pas connexe est dit non connexe, et se décompose en composantes connexes.

Définition A.7.0.8 Composante connexe: Dans un graphe, une composante connexe est un sous-graphe induit maximal connexe. Maximal signifie qu'il n'y a pas de sous-graphe induit connexe plus grand contenant les sommets de la composante.

Définition A.7.0.9 Chaîne eulérienne: Une chaîne eulérienne d'un graphe G est une chaîne qui contient une fois et une seule chaque arête de G .

Définition A.7.0.10 Cycle eulérien: Un cycle eulérien de G est une chaîne eulérienne de G qui est un cycle, c'est-à-dire une chaîne eulérienne dont les extrémités sont confondues.

Définition A.7.0.11 Graphe eulérien: Le graphe G est un graphe eulérien si et seulement si il admet un cycle eulérien.

Théorème 1 : Un graphe connexe G admet un cycle eulérien si et seulement si tous ses sommets sont de degré pair.

Théorème 2 : Un graphe connexe G admet une chaîne eulérienne distincte d'un cycle si et seulement si le nombre de sommets de G de degré impair est égal à 2. Dans ce cas, si n_1 et n_2 sont les deux sommets de G de degré impair, alors le graphe G admet une chaîne eulérienne d'extrémités n_1 et n_2 .

Définition A.7.0.12 Arbre: *Un arbre est un graphe non orienté, acyclique et connexe.*

Définition A.7.0.13 Stable: *Un stable d'un graphe G est un sous-graphe induit de G sans arc. Autrement dit, l'ensemble des sommets d'un stable sont deux à deux non-adjacents dans G .*

Définition A.7.0.14 Multi-ensemble: *On appelle multi-ensemble (aussi dénommé famille d'ensemble) (S_1, \dots, S_N) un ensemble dont les éléments peuvent s'intersecter. En conséquence, $S_i = S_j$ n'impliquera pas $i = j$.*

Un ensemble de partitions d'un même espace forment un multi-ensemble. En effet, deux régions de deux partitions différentes peuvent s'intersecter ou même être égales. Elles seront néanmoins distinguées car elles appartiennent à des partitions différentes.

Définition A.7.0.15 Graphe intersection: *Soit $(\mathcal{P}_1, \dots, \mathcal{P}_T)$ un multi-ensemble ordonné de T partitions du même espace. Le graphe intersection associé est un graphe dont les nœuds sont associés aux régions composant $(\mathcal{P}_1, \dots, \mathcal{P}_T)$, et dont les arcs lient deux nœuds si et seulement si leurs régions ont une intersection non vide.*

Définition A.7.0.16 Adjacence de régions: *Deux régions R_1 et R_2 sont dites adjacentes ssi il existe au moins un pixel p_1 de R_1 et un pixel p_2 de R_2 tel que p_1 et p_2 sont 4-adjacents.*

Définition A.7.0.17 Graphe d'adjacence de régions (RAG): *Le graphe d'adjacence de régions (RAG) de la partition $\mathcal{P} = \{R_i\}_{i=1}^{|\mathcal{P}|}$ est un couple $\mathcal{G} = (\mathcal{N}, \mathcal{E}_s)$, où $\mathcal{N} = \{n_i\}_{i=1}^{|\mathcal{P}|}$ est l'ensemble des nœuds du graphe, et \mathcal{E}_s , l'ensemble des edges spatiaux du graphe. Chaque nœud de \mathcal{N} correspond à une région de \mathcal{P} , et chaque edge spatial correspond à une relation d'adjacence entre régions.*

Bibliographie

- T. Abraham and J. Roddick. Survey of spatio-temporal databases. *GeoInformatica*, 3(1) : 61–99, 1999.
- J. Adams, D. Sabol, V. Kapos, R. Almeida Filho, D. Roberts, M. Smith, and A. Gillespie. Classification of multispectral images based on fractions of endmembers : application to land-cover change in the Brazilian Amazon. *Remote Sensing of Environment*, 52(2) : 137–154, 1995.
- R. Agarwal, C. Aggarwal, and V. Prasad. Depth first generation of long patterns. In *Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 108–118. ACM New York, NY, USA, 2000.
- R. Agrawal and R. Srikant. Fast algorithms for mining association rules. In *Proc. 20th Int. Conf. Very Large Data Bases, VLDB*, volume 1215, pages 487–499, 1994.
- R. Agrawal and R. Srikant. Mining sequential patterns. In *Data Engineering, 1995. Proceedings of the Eleventh International Conference on*, pages 3–14, Mar 1995. doi : 10.1109/ICDE.1995.380415.
- H. Albert-Lorincz. JF (2003) Mining frequent sequential patterns under regular expressions : a highly adaptive strategy for pushing constraints. 3rd SIAM Int. Conf. on Data Mining (SIAM DM’03), USA, 2003.
- J. Allen. Maintaining knowledge about temporal intervals. *Communications of the ACM*, 26(11) :832–843, 1983.
- M. Armstrong. Temporality in spatial databases. *Proceedings of GIS/LIS*, 88(2) :880–9, 1988.
- L. Aurdal, R. Huseby, L. Eikvil, R. Solberg, D. Vikhamar, and A. Solberg. Use of hidden Markov models and phenology for multitemporal satellite image classification : applications to mountain vegetation classification. In *Analysis of Multi-Temporal Remote Sensing Images, 2005 International Workshop on the*, pages 220–224, 2005.
- N. Ayache and O. D. Faugeras. Hyper : a new approach for the recognition and positioning to two-dimensional objects. *IEEE Trans. Pattern Anal. Mach. Intell.*, 8(1) :44–54, 1986. ISSN 0162-8828. doi : <http://dx.doi.org/10.1109/TPAMI.1986.4767751>.
- J. Ayres, J. Flannick, J. Gehrke, and T. Yiu. Sequential pattern mining using a bitmap representation. In *Proceedings of the eighth ACM SIGKDD international conference on knowledge discovery and data mining*, pages 429–435. ACM New York, NY, USA, 2002.
-

- S. Bartalev, F. Achard, D. Erchov, and V. Gond. The potential contribution of SPOT 4/VEGETATION data for mapping Siberian forest cover at continental scale. *Proceedings of Vegetation*, 2 :3–6, 2000.
- Y. Bastide, R. Taouil, N. Pasquier, G. Stumme, and L. Lakhal. Mining frequent patterns with counting inference. *ACM SIGKDD Explorations Newsletter*, 2(2) :66–75, 2000.
- R. Bayardo Jr and R. Agrawal. Mining the most interesting rules. In *Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 145–154. ACM New York, NY, USA, 1999.
- P. S. Beck, C. Atzberger, K. A. Høgda, B. Johansen, and A. K. Skidmore. Improved monitoring of vegetation dynamics at very high latitudes : A new method using modis ndvi. *Remote Sensing of Environment*, 100(3) :321 – 334, 2006. ISSN 0034-4257. doi : DOI:10.1016/j.rse.2005.10.021. URL <http://www.sciencedirect.com/science/article/B6V6V-4HWX9T3-1/2/ea114955feadd4ffd37351a62d3c235>
- A. Belward and C. Valenzuela. *Remote Sensing and Geographical Information Systems for Resource Management in Developing Countries*. Kluwer Academic Print on Demand, 1991.
- S. Beucher and C. Lantuejoul. Use of watersheds in contour detection. In *International Workshop on Image Processing : Real-time Edge and Motion Detection/Estimation, Rennes, France.*, September 1979.
- F. Bodon. A fast apriori implementation. In *FIMI*, 2003.
- E. Bolson, S. Kliman, F. Sheehan, and H. Dodge. Left ventricular segmental wall motion : a new method using local direction information. *Computers in Cardiology*, pages 245–248, 1980.
- J. Boreczky and L. Wilcox. A hidden Markov model framework for video segmentation using audio and image features. In *Acoustics, Speech and Signal Processing, 1998. Proceedings of the 1998 IEEE International Conference on*, volume 6, 1998.
- C. Borgelt. Efficient implementations of apriori and eclat. In *Workshop of Frequent Item Set Mining Implementations*, 2003.
- M. Bosc, F. Heitz, J.-P. Armspach, I. Namer, D. Gounot, and L. Rumbach. Automatic change detection in multimodal serial mri : application to multiple sclerosis lesion evolution. *NeuroImage*, 20(2) :643 – 656, 2003. ISSN 1053-8119. doi : DOI:10.1016/S1053-8119(03)00406-3. URL <http://www.sciencedirect.com/science/article/B6WNP-49MF08F-F/2/8bade0047c38391f0458e2b7a887782>
- P. Bouthemy and E. Francois. Motion segmentation and qualitative dynamic scene analysis from an image sequence. *International Journal of Computer Vision*, 10(2) :157–182, 1993.
- S. Bouzidi, J. Berroir, and I. Herlin. A remote sensing data fusion approach to monitor agricultural areas. *ICPR '98 : Proceedings of the 14th International Conference on Pattern Recognition*, 2 :1387, 1998.

-
- S. Bouzidi, S. Belhaj, I. Herlin, and J. Berroir. An approach for land cover change detection using low spatial resolution data. In *Geoscience and Remote Sensing Symposium, 2003. IGARSS '03. Proceedings. 2003 IEEE International*, volume 3, pages 1585–1587, July 2003.
- F. Bovolo and L. Bruzzone. A multilevel parcel-based approach to change detection in very high resolution multitemporal images. In *Geoscience and Remote Sensing Symposium, 2005. IGARSS '05. Proceedings. 2005 IEEE International*, volume 3, pages 2145–2148, July 2005. doi : 10.1109/IGARSS.2005.1526442.
- F. Bovolo and L. Bruzzone. A theoretical framework for unsupervised change detection based on change vector analysis in the polar domain. *IEEE Transactions on Geoscience and Remote Sensing*, 45(1) :218–236, 2007.
- F. Bovolo, L. Bruzzone, and S. Marchesi. A context-sensitive technique robust to registration noise for change detection in very high resolution multispectral images. In *Geoscience and Remote Sensing Symposium, 2008. IGARSS 2008. IEEE International*, volume 3, pages III –150–III –153, July 2008. doi : 10.1109/IGARSS.2008.4779305.
- Y. Boykov and G. Funka-Lea. Graph cuts and efficient n-d image segmentation. *Int. J. Comput. Vision*, 70(2) :109–131, 2006. ISSN 0920-5691. doi : <http://dx.doi.org/10.1007/s11263-006-7934-5>.
- R. Bremond and F. Marquès. Segmentation-based morphological interpolation of partition sequences. *Mathematical Morphology and its Applications to Image and Signal Processing*, pages 369–376, 1996.
- S. Brin, R. Motwani, J. D. Ullman, and S. Tsur. Dynamic itemset counting and implication rules for market basket data. pages 255–264. ACM Press, 1997.
- L. Bruzzone and D. Prieto. An adaptive parcel-based technique for unsupervised change detection. *International Journal of Remote Sensing*, 21(4) :817–822, 2000.
- L. Bruzzone and S. Serpico. An iterative technique for the detection of land-cover transitions in multitemporal remote-sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 35(4) :858–867, 1997.
- D. Burdick, M. Calimlim, and J. Gehrke. MAFIA : A Maximal Frequent Itemset Algorithm for Transactional Databases. In *PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON DATA ENGINEERING*, pages 443–452. IEEE Computer Society Press; 1998, 2001.
- A. Bykowski and C. Rigotti. A condensed representation to find frequent patterns. In *Symposium on Principles of Database Systems*, 2001. URL citeseer.ist.psu.edu/bykowski01condensed.html.
- P. Cachier and X. Pennec. 3d non-rigid registration by gradient descent on a gaussian-windowed similarity measure using convolutions. In *Mathematical Methods in Biomedical Image Analysis, 2000. Proceedings. IEEE Workshop on*, pages 182–189, 2000. doi : 10.1109/MMBIA.2000.852376.
- T. Calders and B. Goethals. Minimal k-Free Representations of Frequent Sets. *LECTURE NOTES IN COMPUTER SCIENCE*, pages 71–82, 2003.
-

- J. Canny. A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.*, 8(6) :679–698, November 1986. ISSN 0162-8828. URL <http://portal.acm.org/citation.cfm?id=11275>.
- A. Carleer and E. Wolff. Change detection for updates of vector database through region-based classification of VHR satellite data. In *Proceedings of SPIE*, volume 6749, page 674911, 2007.
- P. Ceccato. Operational early warning system using SPOT-VGT and TERRA-MODIS to predict Desert Locust outbreaks. In *Proceedings of the Second International SPOTVEGETATION Users Conference (Editors. F. Veroustraete and E. Bartholomé)*, 2004.
- Chang and Li. Adaptive image region growing. *IEEE Transactions on Image Processing*, 3(6) :868–872, 1994.
- J. Chen, P. Jönsson, M. Tamura, Z. Gu, B. Matsushita, and L. Eklundh. A simple method for reconstructing a high-quality ndvi time-series data set based on the savitzky-golay filter. *Remote Sensing of Environment*, 91(3-4) : 332 – 344, 2004. ISSN 0034-4257. doi : DOI:10.1016/j.rse.2004.03.014. URL <http://www.sciencedirect.com/science/article/B6V6V-4CDS5MR-1/2/3e57e03926adc31538429f34678155f>
- M. Chen, J. Park, and P. Yu. Efficient Data Mining for Path Traversal Patterns. *IEEE Transactions on Knowledge and Data Engineering*, 10(2) :209–221, 1998.
- C. Claramunt and M. Thériault. Managing time in gis : An event-oriented approach. In *Proceedings of the International Workshop on Temporal Databases*, pages 23–42, London, UK, 1995. Springer-Verlag. ISBN 3-540-19945-4.
- P. Clarysse, D. Friboulet, and I. Magnin. Tracking geometrical descriptors on 3-d deformable surfaces : application to the left-ventricular surface of the heart. *Medical Imaging, IEEE Transactions on*, 16(4) :392–404, Aug. 1997. ISSN 0278-0062. doi : 10.1109/42.611349.
- J. Collins and C. Woodcock. An assessment of several linear change detection techniques for mapping forest mortality using multitemporal Landsat TM data. *Remote Sensing of Environment*, 56(1) :66–77, 1996.
- J. B. Collins and C. E. Woodcock. Change detection using the gramm-schmidt transformation applied to mapping forest mortality. *Remote Sensing of Environment*, 50 (3) :267 – 279, 1994. ISSN 0034-4257. doi : DOI:10.1016/0034-4257(94)90076-0. URL <http://www.sciencedirect.com/science/article/B6V6V-488NPVX-7B/2/2b96f9bb530a04e638481b282abab4>
- D. Comaniciu and P. Meer. Robust analysis of feature spaces : color image segmentation. In *Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on*, pages 750–755, Jun 1997. doi : 10.1109/CVPR.1997.609410.
- T. Corpetti, E. Mémin, and P. Pérez. ADAPTATION OF STANDARD OPTIC FLOW METHODS TO FLUID MOTION. In *9th International symposium on flow visualisation*, 2000.
- M. Dalla Mura, J. Benediktsson, F. Bovolo, and L. Bruzzone. An unsupervised technique based on morphological filters for change detection in very high resolution images. *Geoscience and Remote Sensing Letters, IEEE*, 5(3) :433–437, July 2008. ISSN 1545-598X. doi : 10.1109/LGRS.2008.917726.

-
- M. Datcu, K. Seidel, and M. Walessa. Spatial information retrieval from remote-sensing images. i. information theoretical perspective. *Geoscience and Remote Sensing, IEEE Transactions on*, 36(5) :1431–1445, Sep 1998. ISSN 0196-2892. doi : 10.1109/36.718847.
- M. Datcu, H. Daschiel, A. , Pelizzari, M. Quartulli, A. Galoppo, A. Colapicchioni, M. Pastori, K. Seidel, P. Marchetti, and S. D’Elia. Information mining in remote sensing image archives - part a : System concepts. *IEEE Transactions on Geoscience and Remote Sensing*, 41(12), december 2003.
- G. Daughters, E. Alderman, and N. Ingels. A rational approach to the clinical detection of wall motion abnormalities. In *Ventricular Wall Motion, International Symposium Lausanne*, pages 74–82, 1982.
- K. de Beurs and G. Henebry. War, drought, and phenology : changes in the land surface phenology of Afghanistan since 1982. *Journal of Land Use Science*, 3(2) :95–111, 2008.
- K. M. de Beurs and G. M. Henebry. Land surface phenology, climatic variation, and institutional change : Analyzing agricultural land cover change in kazakhstan. *Remote Sensing of Environment*, 89(4) :497 – 509, 2004. ISSN 0034-4257. doi : DOI:10.1016/j.rse.2003.11.006. URL <http://www.sciencedirect.com/science/article/B6V6V-4B9D9BR-4/2/019ab15493818570b48b5f8ff2b2eb5>
- R. De La Briandais. File searching using variable length keys. *AIEE-IRE*, pages 295–298, 1959.
- M. De Martino, F. Causa, and S. Serpico. Classification of optical high resolution images in urban environment using spectral and textural information. In *Geoscience and Remote Sensing Symposium, 2003. IGARSS ’03. Proceedings. 2003 IEEE International*, volume 1, pages 467–469 vol.1, July 2003.
- C. De Roover, A. Herbulot, A. Gouze, E. Debreuve, M. Barlaud, and B. Macq. Multimodal segmentation combining active contours and watersheds. In *Proceedings of the 13th European Signal Processing Conference (EUSIPCO05)*, 2005.
- G. Delyon, F. Galland, and P. Refregier. Minimal stochastic complexity image partitioning with unknown noise model. *Image Processing, IEEE Transactions on*, 15(10) :3207–3212, Oct. 2006. ISSN 1057-7149. doi : 10.1109/TIP.2006.877484.
- A. Demiriz. Asipath : A simple path mining algorithm. In *In Proceedings of the 16th International Conference on Parallel and Distributed Computing and Systems (PDCS 2004)*, pages 9, 2004.
- N. Diehl. Object-oriented motion estimation and segmentation in image sequences. *Signal Processing : Image Communication*, 3(1) :23–56, 1991.
- F. Dufaux and F. Moscheni. Segmentation-based motion estimation for second generation video coding techniques. *Video coding : the second generation approach*, pages 219–263, 1996.
- J. Duncan, P. Shi, A. Amimi, R. Constable, L. Staib, D. Dione, Q. Shi, E. Heller, M. Singer, A. Chakraborty, et al. Toward reliable, non invasive measurement of myocardial function from 4D images. In *Proceedings of SPIE*, volume 2168, page 149, 1994.
-

- B. Dunkel and N. Soparkar. Data organization and access for efficient data mining. In *ICDE*, pages 522–529, 1999. URL citeseer.ist.psu.edu/dunke199data.html.
- P. Eisert and B. Girod. Analyzing facial expressions for virtual conferencing. *IEEE Computer Graphics and Applications*, 18(5) :70–78, 1998.
- D. Filipova-Racheva and M. Hall-Beyer. Smoothing of ndvi time series curves for monitoring of vegetation changes in time. In *Ecological Monitoring and Assessment Network National Science Meeting 2000, Toronto, Ontario, Canada*, January 2000. URL Availableonlineat:http://www.eman-rese.ca/eman/reports/meetings/national2000/toc_posters.html.
- A. Fischer. A simple model for the temporal variations of NDVI at regional scale over agricultural countries. Validation with ground radiometric measurements. *International Journal of Remote Sensing*, 15(7) :1421–1446, 1994.
- A. U. Frank. Qualitative temporal reasoning in gis - ordered time scales. In *advances in GIS research In Proc. of th 6th Int Symp on spatial data handling*, pages 20–35. Taylor and Francis, 1994.
- E. Fredkin. Trie memory. *Communications of the ACM*, 3(9) :490–499, 1960.
- F. Fuchs and H. Le-Men. Efficient subgraph isomorphism with ‘a priori’ knowledge (application to 3d reconstruction of buildings for cartography). In *Proceedings of the Joint IAPR International Workshops on Advances in Pattern Recognition*, pages 427–436, London, UK, 2000. Springer-Verlag. ISBN 3-540-67946-4.
- T. Fung. An assessment of TM imagery for land-cover change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 28(4) :681–684, 1990.
- F. Galland, N. Bertaux, and P. Refregier. Minimum description length synthetic aperture radar image segmentation. *IEEE Transactions on Image Processing*, 12 :995– 1006, 2003.
- F. Galland, N. Bertaux, and P. Réfrégier. Multi-component image segmentation in homogeneous regions based on description length minimization : Application to speckle, poisson and bernoulli noise. *Pattern Recognition*, 38(11) :1926 – 1936, 2005. ISSN 0031-3203. doi : DOI:10.1016/j.patcog.2004.10.002. URL <http://www.sciencedirect.com/science/article/B6V14-4FWV2BK-1/2/07519708b0279399cb531d5e33d4f20>
- A. Galton. Space, time, and the representation of geographical reality. *Topoi*, 20(2) : 173–187, 2001.
- M. Garofalakis, R. Rastogi, and K. Shim. SPIRIT : Sequential pattern mining with regular expression constraints. In *Proceedings of the international conference on very large data bases*, pages 223–234, 1999.
- M. Gelgon and P. Bouthemy. A region-level motion-based graph representation and labeling for tracking a spatial image partition. *Pattern Recognition*, 33(4) :725–740, 2000.
- S. Geman and D. Geman. Stochastic relaxation, gibbs distributions and the bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(6) :721–741, November 1984. doi : 10.1080/02664769300000058. URL <http://dx.doi.org/10.1080/02664769300000058>.

-
- G. Gerig, D. Welte, C. R. Guttmann, A. C. Colchester, and G. Székely. Exploring the discrimination power of the time domain for segmentation and characterization of active lesions in serial mr data. *Medical Image Analysis*, 4(1) :31 – 42, 2000. ISSN 1361-8415. doi : DOI:10.1016/S1361-8415(00)00005-0. URL <http://www.sciencedirect.com/science/article/B6W6Y-40S779D-5/2/97e689eee7acd0005f20f7b4ff8c8de>
- A. Giros. Comparison of partitions of two images for satellite image time series segmentation. In *Geoscience and Remote Sensing Symposium, 2006. IGARSS 2006. IEEE International Conference on*, pages 2592–2595, 31 2006-Aug. 4 2006. doi : 10.1109/IGARSS.2006.670.
- J. Gonzalez, L. Holder, and D. Cook. Application of Graph-Based Concept Learning to the Predictive Toxicology Domain. In *Proceedings of the Predictive Toxicology Challenge Workshop*, 2001.
- K. Gouda and M. Zaki. Efficiently mining maximal frequent itemsets. In *Proceedings of the 2001 IEEE International Conference on Data Mining*, pages 163–170. IEEE Computer Society Washington, DC, USA, 2001.
- K. Gouda, M. Hassaan, and M. Zaki. Prism : A Prime-Encoding Approach for Frequent Sequence Mining. In *Proceedings 7th IEEE International Conference on Data Mining (ICDM 2007, Omaha, NE)*, pages 487–492, 2007.
- S. Goward, C. Tucker, and D. Dye. North American vegetation patterns observed with the NOAA-7 advanced very high resolution radiometer. *Plant Ecology*, 64(1) :3–14, 1985.
- C. Gu and M. Lee. Semantic video object segmentation and tracking using mathematical morphology and perspective motion model. In *IEEE International Conference on Image Processing, Santa Barbara, CA*, volume 10, 1997.
- E. Gudes, S. Shimony, and N. Vanetik. Discovering Frequent Graph Patterns Using Disjoint Paths. *IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING*, pages 1441–1456, 2006.
- L. Gueguen and M. Datcu. Image time-series data mining based on the information-bottleneck principle. *Geoscience and Remote Sensing, IEEE Transactions on*, 45(4) : 827–838, April 2007. ISSN 0196-2892. doi : 10.1109/TGRS.2006.890557.
- L. Gueguen and M. Datcu. A similarity metric for retrieval of compressed objects : Application for mining satellite image time series. *Knowledge and Data Engineering, IEEE Transactions on*, 20(4) :562–575, April 2008. ISSN 1041-4347. doi : 10.1109/TKDE.2007.190718.
- Gueguen, Le Men and Datcu. Analysis of Satellite Image Time Series Based on Information Bottleneck. In *26th Bayesian Inference and Maximum Entropy Methods in Science and Engineering(AIP Conference Proceedings Volume 872)*, volume 872, pages 367–374, 2006.
- L. Guigues. *Modèles multi-échelles pour la segmentation d’images*. Thèse de doctorat, Université de Cergy-Pontoise, 2003.
- J. Guo, J. Kim, and C.-C. J. Kuo. Fast and accurate moving object extraction technique for mpeg-4 object-based video coding. In *Proc. SPIE Visual Comm. and Image Processing*, pages 1210–1221, 1999.
-

- S. Gupta and J. Prince. Stochastic models for div-curl optical flow methods. *IEEE Signal Processing Letters*, 3(2) :32–34, 1996.
- R. Güting, M. Böhlen, M. Erwig, C. Jensen, N. Lorentzos, M. Schneider, and M. Vazirgiannis. A foundation for representing and querying moving objects. *ACM Transactions on Database Systems (TODS)*, 25(1) :1–42, 2000.
- O. Hagolle, G. Dedieu, B. Mougenot, V. Debaecker, B. Duchemin, and A. Meygret. Correction of aerosol effects on multi-temporal images acquired with constant viewing angles : Application to Formosat-2 images. *Remote Sensing of Environment*, 112(4) :1689–1701, 2008.
- A. Hampapur, R. Jain, and T. Weymouth. Production model based digital video segmentation. *Multimedia Tools and Applications*, 1(1) :9–46, 1995.
- J. Han, J. Pei, Y. Yin, and R. Mao. Mining Frequent Patterns without Candidate Generation : A Frequent-Pattern Tree Approach. *Data Mining and Knowledge Discovery*, 8(1) :53–87, 2004.
- A. Hanjalic, R. Lagendijk, and J. Biemond. Automated high-level movie segmentation for advanced video-retrieval systems. *IEEE Transactions on Circuits and Systems for Video Technology*, 9(4) :580–588, 1999.
- Y. Haxhimusa and W. Kropatsch. Segmentation graph hierarchies. *Structural, Syntactic, and Statistical Pattern Recognition*, pages 343–351, 2004. URL <http://www.springerlink.com/content/bnvrlw9697kdynww>.
- P. Heas and M. Datcu. Modeling trajectory of dynamic clusters in image time-series for spatio-temporal reasoning. *IEEE Transactions on Geoscience and Remote Sensing*, 43(7) :1635–1647, 2005.
- P. Heas, E. Memin, N. Papadakis, and A. Szantai. Layered estimation of atmospheric mesoscale dynamics from satellite imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 45(12 Part 2) :4087–4104, 2007.
- C. Hidber. Online association rule mining. *ACM SIGMOD Record*, 28(2) :145–156, 1999.
- J. Hipp, U. Guntzer, and G. Nakhaeizadeh. Mining Association Rules : Deriving a Superior Algorithm by Analyzing Today’s Approaches. *LECTURE NOTES IN COMPUTER SCIENCE*, pages 159–168, 2000.
- J. N. Hird and G. J. McDermid. Noise reduction of ndvi time series : An empirical comparison of selected techniques. *Remote Sensing of Environment*, 113(1) : 248 – 258, 2009. ISSN 0034-4257. doi : DOI:10.1016/j.rse.2008.09.003. URL <http://www.sciencedirect.com/science/article/B6V6V-4TVR2BF-1/2/5a6277f2e3a9f9c17a2a4ba11bf0014>
- D. S. Hochbaum. An efficient algorithm for image segmentation, markov random fields and related problems. *J. ACM*, 48(4) :686–701, 2001. ISSN 0004-5411. doi : <http://doi.acm.org/10.1145/502090.502093>.
- B. Holben. Characteristics of maximum-value composite images from temporal AVHRR data. *International Journal of Remote Sensing*, 7(11) :1417–1434, 1986.

-
- L. Holder, D. Cook, and S. Djoko. Substructure discovery in the subdue system. In *Proc. of the AAAI Workshop on Knowledge Discovery in Databases*, pages 169–180, 1994.
- L. Hong and G. Chen. Segment-based stereo matching using graph cuts. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 1. IEEE Computer Society; 1999, 2004.
- R. Horaud and O. Monga. *Vision par ordinateur : outils fondamentaux*. Hermes, 1995.
- R. Horaud and T. Skordas. Stereo correspondence through feature grouping and maximal cliques. *IEEE Trans. Pattern Anal. Mach. Intell.*, 11(11) :1168–1180, 1989. ISSN 0162-8828. doi : <http://dx.doi.org/10.1109/34.42855>.
- B. Horn and B. Schunck. Determining optical flow. *Computer Vision*, 17 :185–203, 1981.
- S. L. Horowitz and T. Pavlidis. Picture segmentation by a tree traversal algorithm. *J. ACM*, 23(2) :368–388, 1976. ISSN 0004-5411. doi : <http://doi.acm.org/10.1145/321941.321956>.
- H. Horwitz, R. Nalepka, P. Hyde, and J. Morganstern. Estimating the proportion of objects within a single resolution element of a Multispectral Scanner. University of Michigan. *Ann Arbor, Michigan, NASA Contract NAS-9-9784*, 1971.
- A. Huete. A soil-adjusted vegetation index (SAVI). *Remote sensing of environment (USA)*, 1988.
- A. Huete, C. Justice, and W. Van Leeuwen. MODIS vegetation index (MOD13) algorithm theoretical basis document. *Greenbelt : NASA Goddard Space Flight Centre, http://modarch.gsfc.nasa.gov/MODIS/LAND/#vegetation-indices*, 1999.
- P. Héas. *Apprentissage bayésien de structures spatio-temporelles : application à la fouille visuelle des séries temporelles d'images satellites*. Thèse de doctorat, Ecole Nationale Supérieure de l'Aéronautique et de l'Espace (SupAéro), Toulouse, France, avril 2005.
- N. B. Ingels, C. W. Mead, G. T. Daughters, E. B. Stinson, and E. L. Alderman. A new method for assessment of left ventricular segmental wall motion. *Comput. Cardiol.*, page 57–78, 1978.
- J. Inglada, G. Mercier, and T. CNES. A new statistical similarity measure for change detection in multitemporal SAR images and its extension to multiscale change analysis. *IEEE Transactions on Geoscience and Remote Sensing*, 45(5 Part 2) :1432–1445, 2007.
- A. Inokuchi, T. Washio, and H. Motoda. An apriori-based algorithm for mining frequent substructures from graph data. In *Data Mining and Knowledge Discovery*, pages 13–23, 2000.
- ISO/IEC. Information technology - coding of audio-visual objects :visual. Technical report, Doc.ISO/IEC 14496-2 Final Committee Draft, May 1998.
- E. Ito, M. Araki, A. Tani, M. Kanzaki, K. Saret, D. Seila, P. Phearak, L. Sopheap, and P. Sopheavuth. Leaf-shedding phenology in tropical seasonal forests of Cambodia estimated from NOAA satellite images. In *IEEE International Geoscience and Remote Sensing Symposium, 2007. IGARSS 2007*, pages 4331–4335, 2007.
-

S. JHA and N. Unni. Digital change detection of forest conversion of a dry tropical Indian forest region. *International Journal of Remote Sensing*, 15(13) :2543–2552, 1994.

X. Jing, L. Liu, C. Zhang, X. Li, Y. Li, and J. Jia. The extraction of beijing main crops planting area based on time series modis ndvi reconstruction. In *Geoscience and Remote Sensing Symposium, 2005. IGARSS '05. Proceedings. 2005 IEEE International*, volume 4, pages 3020–3023, July 2005.

R. John, J. Chen, N. Lu, K. Guo, C. Liang, Y. Wei, A. Noormets, K. Ma, and X. Han. Predicting plant diversity based on remote sensing products in the semi-arid region of inner mongolia. *Remote Sensing of Environment*, 112(5) : 2018 – 2032, 2008. ISSN 0034-4257. doi : DOI:10.1016/j.rse.2007.09.013. URL <http://www.sciencedirect.com/science/article/B6V6V-4S26RV5-2/2/81cad3dee48310d9c6304f8f6e1997a> Earth Observations for Terrestrial Biodiversity and Ecosystems Special Issue.

P. Jonsson and L. Eklundh. Seasonality extraction by function fitting to time-series of satellite sensor data. *IEEE Transactions on Geoscience and Remote Sensing*, 40(8) : 1824–1832, 2002.

I. Jonyer, D. Cook, and L. Holder. Graph-based hierarchical conceptual clustering. *The Journal of Machine Learning Research*, 2 :19–43, 2002.

C. Jordan. Derivation of leaf-area index from quality of light on the forest floor. *Ecology*, pages 663–666, 1969.

A. Julea, N. Méger, and P. Bolon. On mining METEOSAT and ERS multitemporal images. In *Proc of the 4th conference on Information Mining For Security and Intelligence, ESA-EUSC06*, 2006.

A. Julea, N. Méger, and P. Bolon. On mining pixel based evolution classes in satellite image time series. In *Proc of the 5th conference on Information Mining : pursuing automation of geospatial intelligence for environment and security, ESA-EUSC08*, 2008.

C. Kambhamettu and D. Goldgof. Curvature-based approach to point correspondence recovery in conformal nonrigid motion. *CVGIP. Image understanding*, 60(1) :26–43, 1994.

T. Kanungo, B. Dom, W. Niblack, and D. Steele. A fast algorithm for mdl-based multi-band image segmentation. In *IEEE Conf. CVPR*, pages 609–616, 1994.

M. Kass, A. Witkin, and D. Terzopoulos. Snakes : Active contour models. *International Journal of Computer Vision*, V1(4) :321–331, January 1988. doi : 10.1007/BF00133570. URL <http://dx.doi.org/10.1007/BF00133570>.

Y. Kaufman and D. Tanre. Atmospherically resistant vegetation index (ARVI) for EOS-MODIS. *IEEE Transactions on Geoscience and Remote Sensing*, 30(2) :261–270, 1992.

M. Kawamura, Y. Tsujiko, K. Tsujino, and T. Sakai. Time-series fire-induced forest hazard mapping using Landsat and IKONOS imageries. In *2004 IEEE International Geoscience and Remote Sensing Symposium, 2004. IGARSS'04. Proceedings*, volume 4, 2004.

-
- Q. Ke and T. Kanade. A subspace approach to layer extraction. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 1. IEEE Computer Society; 1999, 2001.
- Z. M. Kedem. Pincer-search : A new algorithm for discovering the maximum frequent set. In *In 6th Intl. Conf. Extending Database Technology*, pages 105–119, 1998.
- A. Klaus, M. Sormann, and K. Karner. Segment-based stereo matching using belief propagation and a self-adapting dissimilarity measure. In *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*, volume 3, pages 15–18, 0-0 2006. doi : 10.1109/ICPR.2006.1033.
- T. Knudsen and B. Olsen. Automated change detection for updates of digital map databases. *Photogrammetric engineering and remote sensing*, 69(11) :1289–1296, 2003.
- G. Koepfler, C. Lopez, and J. M. Morel. A multiscale algorithm for image segmentation by variational method. *SIAM J. Numer. Anal.*, 31(1) :282–299, 1994. ISSN 0036-1429. doi : <http://dx.doi.org/10.1137/0731015>.
- I. Koprinska and S. Carrato. Temporal video segmentation : A survey. *Signal processing : Image communication*, 16(5) :477–500, 2001.
- M. Kryszkiewicz. Concise representation of frequent patterns based on disjunction-free generators. *Data Mining, IEEE International Conference on*, 0 :305, 2001. doi : <http://doi.ieeecomputersociety.org/10.1109/ICDM.2001.989533>.
- M. Kryszkiewicz. Concise Representations of Association Rules. *LECTURE NOTES IN COMPUTER SCIENCE*, pages 92–109, 2002.
- M. Kryszkiewicz and M. Gajek. Why to Apply Generalized Disjunction-Free Generators Representation of Frequent Patterns? *LECTURE NOTES IN COMPUTER SCIENCE*, pages 383–392, 2002.
- M. Kuramochi and G. Karypis. Frequent subgraph discovery. In *Proceedings of the 2001 IEEE International Conference on Data Mining*, pages 313–320, 2001.
- M. Kuramochi and G. Karypis. An Efficient Algorithm for Discovering Frequent Subgraphs. *IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING*, pages 1038–1051, 2004.
- E. Lambin and D. Ehrlich. The surface temperature-vegetation index space for land cover and land-cover change analysis. *International Journal of Remote Sensing*, 17(3) :463–487, 1996.
- G. Langran. *Time in geographic information systems*. Taylor & Francis, 1992.
- G. Langran and N. Chrisman. A framework for temporal geographic information. *Cartographica : The International Journal for Geographic Information and Geovisualization*, 25(3) :1–14, 1988.
- A. Lanterman. Schwarz, Wallace, and Rissanen : Intertwining Themes in Theories of Model Order Estimation. *International Statistical Review*, 69(2) :185–212, 2001.
-

R. Laurini and D. Thompson. Fundamentals of spatial information systems. *San Diego*, 1992.

C. Le Men, A. Julea, M. Méger, N. Datcu, P. Bolon, and H. Maître. Radiometric evolution classification in high resolution satellite images time series. In *Proc of the 5th conference on Information Mining : pursuing automation of geospatial intelligence for environment and security,ESA-EUSC08*, 2008.

Y. G. Leclerc. Constructing simple stable descriptions for image partitioning. *International Journal of Computer Vision*, 3(1) :73–102, 1989.

S. Lee and H. Park. Mining Weighted Frequent Patterns from Path Traversals on Weighted Graph. *IJCSNS*, 7(4) :140, 2007.

R. Leighton, S. Wilt, and R. Lewis. Detection of hypokinesia by a quantitative analysis of left ventricular cineangiograms. *Circulation*, 50(1) :121–127, 1974.

M. Leleu. *Extraction de motifs séquentiels sous contraintes dans des données avec répétitions consécutives*. PhD thesis, INSA de Lyon, 2004.

M. Leleu, C. Rigotti, J. Boulicaut, and G. Euvrard. Constraint-based mining of sequential patterns over datasets with consecutive repetitions. *Lecture notes in computer science*, pages 303–314, 2003.

S. Li. Inexact matching of 3D surfaces. Technical report, technical Report VSSP-TR-3, 1990.

S. Liang. *Quantitative Remote Sensing of Land Surfaces*. Wiley-Interscience, 2005.

Z. Liu, F. Huang, L. Li, and E. Wan. Dynamic monitoring and damage evaluation of flood in north-west Jilin with remote sensing. *International Journal Of Remote Sensing*, 23 (18) :3669–3679, 2002.

D. S. LU, P. MAUSEL, E. S. BRONDIZIO, and MORAN. Change detection of successional and mature forests based on forest stand characteristics using multitemporal TM data in the Altamira, Brazil. In *XXII FIG International Congress, ACSM-ASPRS Annual Conference Proceedings, Washington, DC, USA*, 2002.

M. Ma and F. Veroustraete. Reconstructing pathfinder avhrr land ndvi time-series data for the northwest of china. *Advances in Space Research*, 37(4) : 835 – 840, 2006. ISSN 0273-1177. doi : DOI:10.1016/j.asr.2005.08.037. URL <http://www.sciencedirect.com/science/article/B6V3S-4HBSH4T-4/2/bbaf53def60ca08343a476c8b5fcbe7> Natural Hazards and Oceanographic Processes from Satellite Data.

R. Macleod and R. Congalton. A quantitative comparison of change-detection algorithms for monitoring eelgrass from remotely sensed data. *Photogrammetric Engineering and Remote Sensing*, 64(3) :207–216, 1998.

S. Macomber and C. Woodcock. Mapping and monitoring conifer mortality using remote sensing in the Lake Tahoe Basin. *Remote sensing of environment*, 50(3) :255–266, 1994.

H. Maitre. *Le Traitement des images*. Lavoisier, 2003.

-
- J. Malingreau. Global vegetation dynamics : satellite observations over Asia. *International Journal of Remote Sensing*, 7(9) :1121–1146, 1986.
- J.-F. Mangin. *Mise en correspondance d'images médicales 3D multi-modalités multi-individus pour la corrélation anatomo-fonctionnelle cérébrale*. PhD thesis, ENST Paris, Paris , France, Mar. 1995.
- H. Mannila, H. Toivonen, and A. Verkamo. Discovering frequent episodes in sequences . In *1st Conference on Knowledge Discovery and Data Mining*, pages 210–215, 1995.
- H. Mannila, H. Toivonen, and A. Inkeri Verkamo. Discovery of frequent episodes in event sequences. *Data Mining and Knowledge Discovery*, 1(3) :259–289, 1997.
- I. B. Marcotegui. *Segmentation de séquences d'images en vue du codage*. Thèse de doctorat, Ecole Nationale supérieure des Mines de Paris, Paris, France, 1996.
- F. Marques and C. Molina. Object tracking for content-based functionalities. In *Proceedings of SPIE, the International Society for Optical Engineering*, volume 3024, pages 190–199. Society of Photo-Optical Instrumentation Engineers, 1997.
- F. Marques, B. Llorens, and A. Gasull. Prediction of image partitions using Fourier descriptors : application to segmentation-based coding schemes. *IEEE Transactions on Image Processing*, 7(4) :529–542, 1998.
- F. Masseglia, F. Cathala, and P. Poncelet. The psp approach for mining sequential patterns. In *PKDD '98 : Proceedings of the Second European Symposium on Principles of Data Mining and Knowledge Discovery*, pages 176–184, London, UK, 1998. Springer-Verlag. ISBN 3-540-65068-7.
- A. Matheny and D. Goldgof. The use of three- and four-dimensional surface harmonics for rigid and nonrigid shape recovery and representation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 17(10) :967–981, Oct 1995. ISSN 0162-8828. doi : 10.1109/34.464561.
- R. Mech and M. Wollborn. A noise robust method for 2D shape estimation of moving objects in video sequences considering a moving camera. *Signal Processing*, 66(2) :203–217, 1998.
- T. Meier and K. Ngan. Automatic segmentation of moving objects for video object plane generation. *IEEE Transactions on Circuits and Systems for Video Technology*, 8(5) : 525–538, 1998.
- G. Mercier, L. Hubert-Moy, T. Houet, and P. Gouery. Estimation and monitoring of bare soil/vegetation ratio with SPOT VEGETATION and HRVIR. *IEEE Transactions on Geoscience and Remote Sensing*, 43(2) :348–354, 2005.
- G. Mercier, G. Moser, and S. Serpico. Conditional copula for change detection on heterogeneous sar data. In *Proc of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2007.
- F. Meyer. An overview of morphological segmentation. *International Journal of Pattern Recognition and Artificial Intelligence*, 15(7) :1089–1118, 2001.
-

- F. Meyer and P. Bouthemy. Exploiting the temporal coherence of motion for linking partial spatiotemporal trajectories. In *Computer Vision and Pattern Recognition, 1993. Proceedings CVPR'93., 1993 IEEE Computer Society Conference on*, pages 746–747, 1993.
- F. Meyer and P. Bouthemy. Region-based tracking using affine motion models in long image sequences. *CVGIP Image Understanding*, 60(2) :1994–09, 1994.
- A. Moody and D. M. Johnson. Land-surface phenologies from avhrr using the discrete fourier transform. *Remote Sensing of Environment*, 75(3) :305 – 323, 2001. ISSN 0034-4257. doi : DOI:10.1016/S0034-4257(00)00175-9. URL <http://www.sciencedirect.com/science/article/B6V6V-42TC725-D/2/62e8f0e49a9cfbdb4b9f8b9cd842027>
- F. Moscheni, S. Bhattacharjee, and M. Kunt. Spatio-temporal segmentation based on region merging. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(9) :897–915, 1998.
- D. Muchoney and B. Haack. Change detection for monitoring forest defoliation. *PE&RS-Photogrammetric Engineering & Remote Sensing*, 60(10) :1243–1251, 1994.
- D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Communications on Pure and Applied Mathematics*, 42(5) :577–685, 1989. doi : 10.1002/cpa.3160420503. URL <http://dx.doi.org/10.1002/cpa.3160420503>.
- C. Munyati. Wetland change detection on the Kafue Flats, Zambia, by classification of a multitemporal remote sensing image dataset. *International Journal of Remote Sensing*, 21(9) :1787–1806, 2000.
- A. Nanopoulos and Y. Manolopoulos. Finding Generalized Path Patterns for Web Log Data Mining. *LECTURE NOTES IN COMPUTER SCIENCE*, pages 215–228, 2000.
- A. Neri, S. Colonnese, G. Russo, and P. Talone. Automatic moving object and background separation. *Signal Process.*, 66(2) :219–232, 1998. ISSN 0165-1684. doi : [http://dx.doi.org/10.1016/S0165-1684\(98\)00007-3](http://dx.doi.org/10.1016/S0165-1684(98)00007-3).
- H. T. Nguyen, M. Worring, and R. van den Boomgaard. Watersnakes : Energy-driven watershed segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(3) :330–342, 2003. ISSN 0162-8828. doi : <http://doi.ieeecomputersociety.org/10.1109/TPAMI.2003.1182096>.
- H. Nicolas, S. Pateux, and D. le Guen. Minimum description length criterion and segmentation map coding for region-based video compression. *Circuits and Systems for Video Technology, IEEE Transactions on*, 11(2) :184–198, 2001.
- E. Nishimura, G. Xu, and S. Tsuji. Motion segmentation and correspondence using epipolar constraint. In *Proc. 1st Asian Conf. Computer Vision*, pages 199–204, 1993.
- J. Ohm. Core experiments on multifunctional and advanced layered coding aspects of MPEG-4 video. *Doc. ISO/IEC JTC1/SC29/WG11*, 2176, may 1998.

-
- S. I. Olsen. Epipolar line estimation. In *ECCV '92 : Proceedings of the Second European Conference on Computer Vision*, pages 307–311, London, UK, 1992. Springer-Verlag. ISBN 3-540-55426-2.
- L. Olsson and L. Eklundh. Fourier series for analysis of temporal sequences of satellite sensor imagery. *International Journal of Remote Sensing*, 15(18) :3735–3741, 1994.
- S. Osher and J. A. Sethian. Fronts propagating with curvature-dependent speed : Algorithms based on hamilton-Jacobi formulations. *Journal of Computational Physics*, 79 :12–49, 1988. URL <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.46.1266>.
- N. Paragios and R. Deriche. Geodesic active regions and level set methods for supervised texture segmentation. *Int. J. Comput. Vision*, 46(3) :223–247, 2002. ISSN 0920-5691. doi : <http://dx.doi.org/10.1023/A:1014080923068>.
- J. Park, M. Chen, and P. Yu. An effective hash-based algorithm for mining association rules. In *Proc. ACM-SIGMOD Int. Conf. Management of Data (SIGMOD'95), San Jose*, pages 175–186, 1995.
- J. Park, D. Metaxas, and L. Axel. Analysis of left ventricular wall motion based on volumetric deformable models and MRI-SPAMM. *Medical Image Analysis*, 1(1) :53–71, 1996a.
- J. Park, D. Metaxas, A. Young, and L. Axel. Deformable models with parameter functions for cardiac motion analysis from tagged MRI data. *IEEE Transactions on Medical Imaging*, 15(3) :278–289, 1996b.
- N. Pasquier, Y. Bastide, R. Taouil, and L. Lakhal. Discovering frequent closed itemsets for association rules. In *Database Theory-ICDT'99*, pages 398–416, 1999.
- J. Pei, J. Han, and R. Mao. Closet : An efficient algorithm for mining frequent closed itemsets. In *proc. SIGMOD Int'l Workshop Data Mining and knowledge Discovery*, pages 21–30, 2000.
- J. Pei, J. Han, B. Mortazavi-Asl, H. Pinto, Q. Chen, U. Dayal, and M. Hsu. PrefixSpan : Mining Sequential Patterns Efficiently by Prefix-Projected Pattern. In *IEEE Int. Conference on Data Engineering*, 2001.
- A. Peng and W. Pieczynski. Adaptive mixture estimation and unsupervised local bayesian image segmentation. *Graphical Models and Image Processing*, 57(5) :389 – 399, 1995. ISSN 1077-3169. doi : DOI:10.1006/gmip.1995.1033. URL <http://www.sciencedirect.com/science/article/B6WG4-45N1R3X-D/2/4706982cad42f5acbd86f105654d748>
- N. Pettorelli, J. O. Vik, A. Mysterud, J.-M. Gaillard, C. J. Tucker, and N. C. Stenseth. Using the satellite-derived ndvi to assess ecological responses to environmental change. *Trends in Ecology & Evolution*, 20(9) :503 – 510, 2005. ISSN 0169-5347. doi : DOI:10.1016/j.tree.2005.05.011. URL <http://www.sciencedirect.com/science/article/B6VJ1-4GC1R3X-2/2/92dc445fbaa1aed02419ba5dc68f1d2>
- D. Peuquet. It's about time : A conceptual framework for the representation of temporal dynamics in geographic information systems. *Annals of the Association of American Geographers*, pages 441–461, 1994.
-

- D. Peuquet and N. Duan. An event-based spatiotemporal data model (ESTDM) for temporal analysis of geographical data. *International journal of geographical information science*, 9(1) :7–24, 1995.
- J. Piwowar, D. Peddle, and D. Sauchyn. Identifying Ecological Variability in Vegetation Dynamics through Temporal Mixture Analysis. In *IEEE International Conference on Geoscience and Remote Sensing Symposium, 2006. IGARSS 2006*, pages 3766–3769, 2006.
- A. Prakash and R. Gupta. Land-use mapping and change detection in a coal mining area—a case study in the Jharia coalfield, India. *International journal of remote sensing*, 19(3) : 391–410, 1998.
- J. Qi, A. Chehbouni, A. Huete, Y. Kerr, and S. Sorooshian. A modified soil adjusted vegetation index. *Remote Sensing of Environment*, 48(2) :119–126, 1994.
- A. Raza, W. Kainz, and R. Sliuzas. Design and Implementation of a Temporal GIS For Monitoring Urban Land Use. *Change*, pages 41742–7, 1996.
- A. Raza, W. Kainz, and R. Sliuzas. Design and Implementation of a Temporal GIS for Monitoring the Urban Land Use Change. *Proceedings of the Spatial Information Technology Towards*, pages 417–427, 2000.
- A. Renolen. History graphs : Conceptual modelling of spatiotemporal data. *Proceedings of GIS Frontiers in Business and Science*, 1996.
- E. Reusens. Joint optimization of representation model and frame segmentation for generic video compression. *Signal Processing*, 46(1) :105–117, 1995. doi : 10.1016/0165-1684(95)00075-O.
- D. Rey, G. Subsol, H. Delingette, and N. Ayache. Automatic detection and segmentation of evolving processes in 3d medical images : Application to multiple sclerosis. *Medical Image Analysis*, 6(2) :163 – 179, 2002. ISSN 1361-8415. doi : DOI:10.1016/S1361-8415(02)00056-7. URL <http://www.sciencedirect.com/science/article/B6W6Y-45H930X-1/2/a7898fa574749d88ed7df61d19ddf75>
- M. Ridd and J. LIU. A comparison of four algorithms for change detection in an urban environment. *Remote sensing of environment*, 63(2) :95–100, 1998.
- E. Rignot and J. van Zyl. Change detection techniques for ers-1 sar data. *Geoscience and Remote Sensing, IEEE Transactions on*, 31(4) :896–906, Jul 1993. ISSN 0196-2892. doi : 10.1109/36.239913.
- J. Rissanen. Modeling by shortest data description. *Automatica*, 14(5) :465–471, 1978.
- J. Rissanen. A universal prior for integers and estimation by minimum description length. *The Annals of Statistics*, 11(2) :416–431, 1983. ISSN 00905364. URL <http://www.jstor.org/stable/2240558>.
- J. Rissanen. *Stochastic Complexity in Statistical Inquiry Theory*. World Scientific Publishing Co., Inc. River Edge, NJ, USA, 1989.

-
- A. Robin, S. Le Hégarat-Masclé, and L. Moisan. Unsupervised subpixelic classification using coarse-resolution time series and structural information. *Geoscience and Remote Sensing, IEEE Transactions on*, 46(5) :1359–1374, May 2008. ISSN 0196-2892. doi : 10.1109/TGRS.2008.916477.
- A. Robin, L. Moisan, and S. Le Hégarat-Masclé. An a-contrario approach for sub-pixel change detection in satellite imagery. 2009. URL <http://hal.archives-ouvertes.fr/hal-00399698/en/>.
- A. Rosenfeld, R. A. Hummel, and S. W. Zucker. Scene labeling by relaxation operations. *Systems, Man and Cybernetics, IEEE Transactions on*, 6(6) :420–433, June 1976. ISSN 0018-9472. doi : 10.1109/TSMC.1976.4309519.
- P. Salembier, F. Marques, M. Pardas, J. Morros, I. Corset, S. Jeannin, L. Bouchard, F. Meyer, and B. Marcotegui. Segmentation-based video coding system allowing the manipulation of objects. *Circuits and Systems for Video Technology, IEEE Transactions on*, 7(1) :60–74, Feb 1997. ISSN 1051-8215. doi : 10.1109/76.554418.
- A. Savasere, E. Omiecinski, and S. Navathe. An Efficient Algorithm for Mining Association Rules in Large Databases. In *Int. Conf. Very Large Data Bases (VLDB'95), Zurich, Switzerland*, pages 432–443. Georgia Institute of Technology, 1995.
- M. Schroder, H. Rehrauer, K. Seidel, and M. Datcu. Spatial information retrieval from remote-sensing images. ii. gibbs-markov random fields. *Geoscience and Remote Sensing, IEEE Transactions on*, 36(5) :1446–1455, Sep 1998. ISSN 0196-2892. doi : 10.1109/36.718848.
- J. Settle and N. Drake. Linear mixing and the estimation of ground cover proportions. *International Journal of Remote Sensing*, 14(6) :1159–1177, 1993.
- L. S. Shapiro and M. Brady. Rejecting Outliers and Estimating Errors in an Orthogonal-Regression Framework. *Royal Society of London Philosophical Transactions Series A*, 350 :407–439, Mar. 1995.
- P. Shenoy, J. Haritsa, S. Sudarshan, G. Bhalotia, M. Bawa, and D. Shah. Turbo-charging vertical mining of large databases. *ACM SIGMOD Record*, 29(2) :22–33, 2000.
- J. Shi and J. Malik. Normalized cuts and image segmentation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(8) :888–905, Aug 2000. ISSN 0162-8828. doi : 10.1109/34.868688.
- A. Singh. Change detection in the tropical forest environment of northeastern India using Landsat. *Remote sensing and tropical land management*, pages 237–254, 1986.
- P. Smits and A. Annoni. Toward specification-driven change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 38(3) :1484–1488, 2000.
- R. T. Snodgrass. Theories and methods of spatio-temporal reasoning in geographic space. In *Temporal databases*, pages 22–64, 1992. doi : 10.1007/3-540-55966-3_2.
- M. Spanner, L. PIERCE, S. RUNNING, and D. PETERSON. The seasonality of AVHRR data of temperate coniferous forests : Relationship with leaf area index. *Remote Sensing of Environment*, 33(2) :97–112, 1990.
-

- R. Srikant and R. Agrawal. Mining sequential patterns : Generalizations and performance improvements. In *Advances in Database Technology—EDBT'96 : 5th International Conference on Extending Database Technology, Avignon, France, March 25-29, 1996 : Proceedings*, page 3. Springer, 1996.
- D. Stewart, H. Dodge, and M. Frimer. Quantitative analysis of regional myocardial performance in coronary artery disease. *Cardiovascular imaging and image processing : Theory and practice- 1975*, pages 217–224, 1975.
- D. Suter. Motion estimation and vector splines. In *1994 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1994. Proceedings CVPR'94.*, pages 939–942, 1994.
- D. Swets, B. Reed, J. Rowland, and S. Marko. A weighted least-squares approach to temporal NDVI smoothing. In *Proceedings of the 1999 ASPRS Annual Conference : From Image to Information, Portland, Oregon, 1999*.
- A. Tait and X. Zheng. Mapping frost occurrence using satellite data. *Journal of Applied Meteorology*, 42(2) :193–203, 2003.
- J. Thirion. Image matching as a diffusion process : an analogy with Maxwell's demons. *Medical Image Analysis*, 2(3) :243–260, 1998.
- H. Toivonen et al. Sampling large databases for association rules. In *Proceedings of the International Conference on Very Large Data Bases*, pages 134–145. INSTITUTE OF ELECTRICAL & ELECTRONICS ENGINEERS (IEEE), 1996.
- C. Tucker, C. Vanpraet, E. Boerwinkel, and A. Gaston. Satellite remote sensing of total dry matter production in the Senegalese Sahel. *Remote Sensing of Environment*, 1983.
- A. Van Dijk, S. Callis, C. Sakamoto, and W. Decker. Smoothing vegetation index profiles : an alternative method for reducing radiometric disturbance in NOAA/AVHRR data. *Photogrammetric engineering and remote sensing (USA)*, 1987.
- P. F. Velleman. Definition and comparison of robust nonlinear data smoothing algorithms. *Journal of the American Statistical Association*, 75(371) :609–615, 1980. ISSN 01621459. URL <http://www.jstor.org/stable/2287657>.
- W. Verhoef, M. Menenti, and S. Azzali. A colour composite of NOAA-AVHRR-NDVI based on time series analysis (1981-1992). *International Journal of Remote Sensing*, 17(2) :231–235, 1996.
- L. A. Vese and T. F. Chan. A multiphase level set framework for image segmentation using the mumford and shah model. *International Journal of Computer Vision*, 50(3) :271–293, December 2002. doi : 10.1023/A:1020874308076. URL <http://dx.doi.org/10.1023/A:1020874308076>.
- L. Vincent and P. Soille. Watersheds in digital spaces : an efficient algorithm based on immersion simulations. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 13(6) :583–598, 1991.

-
- M. Wachowicz and R. Healey. Towards temporality in GIS. In *Innovations in Gis : Selected Papers from the First National Conference on Gis Research Uk*, page 105. Taylor & Francis, 1994.
- C. Wallace and P. Freeman. Single-factor analysis by minimum message length estimation. *Journal of the Royal Statistical Society, B*, 54(1) :195–209, 1992.
- V. Walter. Object-based classification of remote sensing data for change detection. *ISPRS Journal of Photogrammetry and Remote Sensing*, 58(3-4) :225 – 238, 2004. ISSN 0924-2716. doi : DOI:10.1016/j.isprsjprs.2003.09.007. URL <http://www.sciencedirect.com/science/article/B6VF4-4B6KH0Y-3/2/e0d5c8e54a2a10e1a9bdd34ba351266>
- Integration of Geodata and Imagery for Automated Refinement and Update of Spatial Databases.
- D. Wang. Unsupervised video segmentation based on watersheds and temporaltracking. *IEEE Transactions on Circuits and Systems for Video Technology*, 8(5) :539–546, 1998.
- F. Wang. A knowledge-based vision system for detecting land changes at urbanfringes. *IEEE Transactions on Geoscience and Remote Sensing*, 31(1) :136–145, 1993.
- J. Wang and E. Adelson. Representing moving images with layers. *IEEE Transactions on Image Processing*, 3(5) :625–638, 1994.
- J. Wang, J. Han, and J. Pei. CLOSET+ : searching for the best strategies for mining frequent closed itemsets. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 236–245. ACM New York, NY, USA, 2003a.
- Q. Wang, M. Watanabe, S. Hayashi, and S. Murakami. Using NOAA AVHRR data to assess flood damage in China. *Environmental monitoring and assessment*, 82(2) :119–148, 2003b.
- R. Weismiller, S. Kristof, D. Scholz, P. Anuta, S. Momin, and I. Purdue Univ., Lafayette. Change detection in coastal zone environments(by Landsat MSS data analysis). *Photogrammetric Engineering and Remote Sensing*, 43 :1533–1539, 1977.
- M. Worboys. A model for spatio-temporal information. In *Proceedings of the 5th International Symposium on Spatial Data Handling*, volume 2, pages 602–611, 1992.
- M. Worboys. A unified model for spatial and temporal information. *The Computer Journal*, 37(1) :26–34, 1994.
- M. Worboys, H. Hearnshaw, and D. Maguire. Object-oriented data modelling for spatial databases. *International Journal of Geographical Information Science*, 4(4) :369–383, 1990.
- Z. Wu and R. Leahy. An optimal graph theoretic approach to data clustering : theory and its application to image segmentation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 15(11) :1101–1113, Nov 1993. ISSN 0162-8828. doi : 10.1109/34.244673.
- J. Xiao and M. Shah. Motion layer extraction in the presence of occlusion using graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(10) :1644–1659, 2005.
-

- G. Xu, E. Nishimura, and S. Tsuji. Image correspondence and segmentation by epipolar lines : Theory, algorithm and applications. Technical report, Technical report, Dept. of Systems Engineering, Osaka University, Japan, 1993.
- Y.Lemur. Segmentation multitemporelle d'une séquence d'images spot, *MsC thesis*. Master's thesis, INPG, september 2004.
- M. Yuan. Use of a Three-Domain Representation to Enhance GIS Support for Complex Spatiotemporal Queries. *Transactions in GIS*, 3(2) :137–159, 1999.
- R. Zabih, J. Miller, and K. Mai. A feature-based algorithm for detecting and classifying scene breaks. In *Proceedings of the third ACM international conference on Multimedia*, pages 189–200. ACM New York, NY, USA, 1995.
- M. Zaki. Sequence mining in categorical domains : incorporating constraints. In *Proceedings of the ninth international conference on Information and knowledge management*, pages 422–429. ACM New York, NY, USA, 2000a.
- M. Zaki. SPADE : An Efficient Algorithm for Mining Frequent Sequences. *Machine Learning*, 42(1) :31–60, 2001.
- M. Zaki. Efficient enumeration of frequent sequences. In *Proceedings of the seventh international conference on Information and knowledge management*, pages 68–75. ACM New York, NY, USA, 1998.
- M. Zaki. Generating non-redundant association rules. In *Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 34–43. ACM New York, NY, USA, 2000b.
- M. Zaki and K. Gouda. Fast vertical mining using diffsets. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 326–335. ACM New York, NY, USA, 2003.
- M. Zaki, S. Parthasarathy, M. Ogihara, W. Li, et al. New algorithms for fast discovery of association rules. In *3rd Intl. Conf. on Knowledge Discovery and Data Mining*, volume 20, 1997.
- M. J. Zaki. Scalable algorithms for association mining. *IEEE Transactions on Knowledge and Data Engineering*, 12 :372–390, 2000c.
- H. Zhang, A. Kankanhalli, and S. Smoliar. Automatic partitioning of full-motion video. *Readings in Multimedia Computing and Networking*, 1993.
- X. Zhang, M. Friedl, C. Schaaf, A. Strahler, J. Hodges, F. Gao, B. Reed, and A. Huete. Monitoring vegetation phenology using MODIS. *Remote Sensing of Environment*, 84 (3) :471–475, 2003.
- Z. Zhang, R. Deriche, O. Faugeras, and Q.-T. Luong. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. *Artif. Intell.*, 78(1-2) :87–119, 1995. ISSN 0004-3702. doi : [http://dx.doi.org/10.1016/0004-3702\(95\)00022-4](http://dx.doi.org/10.1016/0004-3702(95)00022-4).

S. C. Zhu, T. S. Lee, and A. L. Yuille. Region competition : Unifying snakes, region growing, energy/bayes/MDL for multi-band image segmentation. In *ICCV*, pages 416–, 1995. URL citeseer.ist.psu.edu/zhu95region.html.

Publications

- C. Le Men, H.Maître, M.Datcu, Spatio-Temporal Segmentation of Satellite Image Time Series, In *CNES-DLR -CCT workshop*, Toulouse, Oct 2008
- C. Le Men, A. Julea, N. Méger, M. Datcu, P.Bolon, H.Maître, Radiometric evolution classification in high resolution Satellite Image Time Series (STIS), In *ESA-EUSC on Image Information Mining : pursuing automation of geospatial intelligence for environment and security 2008*, Frascati, Italy, March 2008
- C. Le Men, H. Maître, M. Datcu, Minimum Description Length principle applied to the segmentation of High Resolution Satellite Image Time Serie, In *ESA-EUSC on Image Information Mining : pursuing automation of geospatial intelligence for environment and security 2008*, Frascati, Italy, March 2008
- C. Le Men, M.Datcu, Hierarchical concept for reasoning on segmentations of Satellite Image Time Series : first result, In *ESA-EUSC on Image Information Mining for Security and Intelligence*, Torrejon air base - Madrid (Spain), Nov 2006
- L. Gueguen, C. Le Men and M. Datcu, Analysis of Satellite Image Time Series based on Information Bottleneck In *26th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, Paris, July 2006

Glossaire

couverture La couverture K d'un motif m est l'ensemble des identifiants des objets couvrants de m : $K(f) = \{j \in \{1, \dots, |\mathcal{O}|\} | f \subseteq o_j\}$. 102, 106, 217

couvrant Un objet ou un motif couvre m_k si $m \subseteq o$, respectivement $m_k \subset m_{k+1}$. On dit que m_{k+1} est un motif couvrant de m_k . 102, 103

fermé Un motif fréquent est fermé si il n'existe pas de motif couvrant de même fréquence. L'ensemble des motifs fréquents fermés est noté C . 103, 217

fréquence La fréquence σ d'un k -motif m dans la base de données est le nombre d'objets qui contiennent m . Dans la littérature, le terme support est aussi fréquemment utilisé pour désigner la fréquence.. 102, 217

fréquent Un k -motif est fréquent si il est contenu dans au moins σ_{\min} objets, où σ_{\min} est un paramètre de l'algorithme.. 102, 217

identifiant Les identifiants sont les indices associés à chacun des objets.. 102

instance Une instance d'un objet spatio-temporel est son état à un instant donné.. 139

k-motif Motif de longueur k . 102

libre Un motif fréquent m est libre si $\sigma(m) > \sigma_{\min}$. L'ensemble des motifs fréquents libres est noté L . 103, 217

longueur La longueur d'un motif désigne son cardinal : le nombre de symboles qu'il contient.. 102

maximal Un motif fréquent est maximal si il n'existe pas de motif couvrant de fréquence supérieure. L'ensemble des motifs fréquents maximal, noté M est parfois aussi appelé bordure positive.. 103, 217

motif Un motif désigne un ensemble de symboles. 102

n-extension Une n -extension de m est la concaténation du motif m avec un n -motif. 102

objet Chaque objet o_i contient un ensemble de k symboles $s_i \in \mathcal{S}$. 102

Acronyms

ADAM STIS acquise par la constellation SPOT (1,2, et 4) dans le cadre du projet d'Assimilation de Données pour l'Agro-Modélisation.. 21, 23, 42, 44, 49

Apriori Ce terme fait référence à la propriété Apriori, et à l'algorithme de recherche de motifs fréquents qui en découle. La propriété stipule qu'aucun motif couvrant d'un motif non-fréquent n'est fréquent. Voir A.3.1 pour une description de l'algorithme.. 101, 103, 104

ECLAT Algorithme de recherche de motifs fréquents dans des bases de données utilisant un parcours en profondeur de l'arbre de préfixe représentant les motifs fréquents et une représentation verticale de la base de données.. 101, 104, 106–108, 144, 146, 148, 180

FIFO structure de stockage traitée de sorte que les premiers éléments entrés dans la file sont les premiers utilisés.. 95

FP-Growth TODO. 101

image-I image codée en intra. 34

image-P image codée par rapport à la précédente image. 34

LDSOSTAG graphe étiqueté et dirigé d'adjacence spatio-temporelle de régions. 142

LPE algorithme de segmentation par Ligne de Partage des Eaux. 33, 34, 43, 92, 93

MDL Principe de minimisation de la longueur de description. 23, 92–94, 179

MT Multi-temporel first. 140

NDVI Indice de végétation par différence normalisée, Normalized Difference Vegetation Index.. 38, 39, 81

SO Objet spatial. 22–24, 79–81, 84, 85, 87, 115, 116, 139, 140, 143, 150, 151

SOTAG graphe d'adjacence spatio-temporelle de régions. 116, 117, 125, 131

SOTAG Graphe d'adjacence temporelle d'objets spatiaux.. 22–24, 85–87, 115, 116, 134, 137, 142, 179

STIS Séquence temporelle d'images satellitaires. Ce terme fait généralement référence aux STIS à haute résolution, STIS-HR first. 21–23, 30, 31, 37, 39, 41–44, 50, 51, 53, 83–88, 98, 115–119, 121, 125–127, 131, 140, 142, 143, 152, 179–182, 215, 216

STIS-CR Séquence temporelle d'images satellitaires à haute résolution à résolution grossière. 38, 41, 42, 50, 83, 88

- STIS-HR** Séquence temporelle d'images satellitaires à haute résolution à haute résolution.
On ne précise HR que pour insister sur la spécificité de la haute résolution. 21, 22, 29, 42, 43, 50, 88, 179, 215
- STISc** STIS composites. Chaque image est obtenue à partir de plusieurs images de façon à supprimer les pixels de nuage.. 39–41
- STO** Objet spatio-temporel.. 22, 23, 80–88, 115, 116, 139–143, 146, 147, 149–152
- TL** Temporellement localisé first. 140–142, 149–152

Notations

- \mathcal{C} Contour d'une région.. 95
- d Dimension de l'espace des caractéristiques. 96
- D Domaine spatial de l'image. 94
- \mathcal{S} Ensemble non ordonné de symboles. 101, 102
- \mathcal{SG} Graphe d'adjacence de régions spatiales. 98
- \mathcal{ASG} Graphe d'adjacence de régions spatiales attribué. 98
- Im Image observée. 94
- \mathbf{Im}_{R_i} sous-image : restriction de l'image Im à la région spatiale R_i . 97
- \mathcal{L}_{fid} Longueur de description de l'erreur de description de l'image par son modèle. 95
- \mathcal{L}_{geom} Longueur de description géométrique. 94
- \mathcal{L}_{noise} Longueur de description du bruit. 94
- \mathcal{L}_{radiom} Longueur de description radiométrique. 94
- Γ Matrice de covariance. 97
- \tilde{Im} Modèle de l'image. 94, 96
- μ Moyenne. 97
- \mathcal{P} Partition du domaine de l'image. 94
- R Région spatiale. 97
- $N_x * N_y$ Taille de l'image. N_x : nombre de colonnes, N_y : nombre de lignes. 95, 96
- \mathcal{C} Ensemble des motifs fréquents fermés. 103
- \mathcal{F} Ensemble des motifs fréquents. 102
- \mathcal{F} Ensemble des k-motifs fréquents. 102
- \mathcal{L} Ensemble des motifs fréquents libres. 103
- \mathcal{M} Ensemble des motifs fréquents maximaux. 103
- K couverture. 102
- s Symbole. Par exemple un type d'article, un type de nucléotide dans une séquence ADN ou encore un vecteur de quantification dans l'espace des caractéristiques quantifié des objets spatiaux de la scène : un centroïde de la quantification des radiométries des régions de la scène.. 101
- m Motif c'est-à-dire un ensemble de symboles.. 102, 213
- σ fréquence. 102

