Classification non supervisée de données spatio-temporelles multidimensionnelles : Applications à l'imagerie

Simon Mure

N°d'ordre NNT : 2016LYSEI130

THESE de DOCTORAT DE L'UNIVERSITE DE LYON
opérée au sein de
INSA LYON

Ecole Doctorale ED 160
Electronique, Electrotechnique, Automatique

Spécialité de doctorat : Traitement du signal et de l'image

Soutenue publiquement le 02/12/2016, par :
Simon Mure

Classification non supervisée de données spatio-temporelles multidimensionnelles.

Applications à l'imagerie.

Devant le jury composé de :

DUCOTTET Christophe	Professeur, Université de Saint Etienne	Président
DOUZAL Ahlame	MDC, HDR, Université Joseph Fourier	Rapporteure
MACAIRE Ludovic	Professeur, Université de Lille	Rapporteur
MARTEAU Pierre-François	Professeur, Université Bretagne Sud	Examinateur
GUTTMANN Charles	Professeur, Harvard Medical School	Membre invité
GRENIER Thomas	MDC, INSA Lyon	Co-directeur de thèse
BENOIT-CATTIN Hugues	Professeur, INSA Lyon	Directeur de thèse

Département FEDORA – INSA Lyon - Ecoles Doctorales – Quinquennal 2016-2020

SIGLE	ECOLE DOCTORALE	NOM ET COORDONNEES DU RESPONSABLE
CHIMIE	**CHIMIE DE LYON** **http://www.edchimie-lyon.fr** Sec : Renée EL MELHEM Bat Blaise Pascal 3e etage secretariat@edchimie-lyon.fr Insa : R. GOURDON	M. Stéphane DANIELE Institut de Recherches sur la Catalyse et l'Environnement de Lyon IRCELYON-UMR 5256 Équipe CDFA 2 avenue Albert Einstein 69626 Villeurbanne cedex directeur@edchimie-lyon.fr
E.E.A.	**ELECTRONIQUE, ELECTROTECHNIQUE, AUTOMATIQUE** **http://edeea.ec-lyon.fr** Sec : M.C. HAVGOUDOUKIAN Ecole-Doctorale.eea@ec-lyon.fr	M. Gérard SCORLETTI Ecole Centrale de Lyon 36 avenue Guy de Collongue 69134 ECULLY Tél : 04.72.18 60.97 Fax : 04 78 43 37 17 Gerard.scorletti@ec-lyon.fr
E2M2	**EVOLUTION, ECOSYSTEME, MICROBIOLOGIE, MODELISATION** **http://e2m2.universite-lyon.fr** Sec : Sylvie ROBERJOT Bât Atrium - UCB Lyon 1 04.72.44.83.62 Insa : H. CHARLES secretariat.e2m2@univ-lyon1.fr	M. Fabrice CORDEY CNRS UMR 5276 Lab. de géologie de Lyon Université Claude Bernard Lyon 1 Bât Géode 2 rue Raphaël Dubois 69622 VILLEURBANNE Cédex Tél : 06.07.53.89.13 cordey@univ-lyon1.fr
EDISS	**INTERDISCIPLINAIRE SCIENCES-SANTE** **http://www.ediss-lyon.fr** Sec : Sylvie ROBERJOT Bât Atrium - UCB Lyon 1 04.72.44.83.62 Insa : M. LAGARDE secretariat.ediss@univ-lyon1.fr	Mme Emmanuelle CANET-SOULAS INSERM U1060, CarMeN lab, Univ. Lyon 1 Bâtiment IMBL 11 avenue Jean Capelle INSA de Lyon 696621 Villeurbanne Tél : 04.72.68.49.09 Fax :04 72 68 49 16 Emmanuelle.canet@univ-lyon1.fr
INFOMATHS	**INFORMATIQUE ET MATHEMATIQUES** **http://infomaths.univ-lyon1.fr** Sec :Renée EL MELHEM Bat Blaise Pascal 3e etage infomaths@univ-lyon1.fr	Mme Sylvie CALABRETTO LIRIS – INSA de Lyon Bat Blaise Pascal 7 avenue Jean Capelle 69622 VILLEURBANNE Cedex Tél : 04.72. 43. 80. 46 Fax 04 72 43 16 87 Sylvie.calabretto@insa-lyon.fr
Matériaux	**MATERIAUX DE LYON** **http://ed34.universite-lyon.fr** Sec : M. LABOUNE PM : 71.70 –Fax : 87.12 Bat. Direction Ed.materiaux@insa-lyon.fr	M. Jean-Yves BUFFIERE INSA de Lyon MATEIS Bâtiment Saint Exupéry 7 avenue Jean Capelle 69621 VILLEURBANNE Cedex Tél : 04.72.43 71.70 Fax 04 72 43 85 28 jean-yves.buffiere@insa-lyon.fr
MEGA	**MECANIQUE, ENERGETIQUE, GENIE CIVIL, ACOUSTIQUE** **http://mega.universite-lyon.fr** Sec : M. LABOUNE PM : 71.70 –Fax : 87.12 Bat. Direction mega@insa-lyon.fr	M. Philippe BOISSE INSA de Lyon Laboratoire LAMCOS Bâtiment Jacquard 25 bis avenue Jean Capelle 69621 VILLEURBANNE Cedex Tél : 04.72 .43.71.70 Fax : 04 72 43 72 37 Philippe.boisse@insa-lyon.fr
ScSo	**ScSo*** **http://recherche.univ-lyon2.fr/scso/** Sec : Viviane POLSINELLI Brigitte DUBOIS Insa : J.Y. TOUSSAINT Tél : 04 78 69 72 76 viviane.polsinelli@univ-lyon2.fr	M. Christian MONTES Université Lyon 2 86 rue Pasteur 69365 LYON Cedex 07 Christian.montes@univ-lyon2.fr

*ScSo : Histoire, Géographie, Aménagement, Urbanisme, Archéologie, Science politique, Sociologie, Anthropologie

Résumé

Avec l'augmentation considérable d'acquisitions de données temporelles dans les dernières décennies comme les systèmes GPS, les séquences vidéo ou les suivis médicaux de pathologies ; le besoin en algorithmes de traitement et d'analyse efficaces d'acquisition longitudinales n'a fait qu'augmenter. Dans cette thèse, nous proposons une extension du formalisme mean-shift, classiquement utilisé en traitement d'images, pour le groupement de séries temporelles multidimensionnelles. Nous proposons aussi un algorithme de groupement hiérarchique des séries temporelles basé sur la mesure de dynamic time warping afin de prendre en compte les déphasages temporels. Ces choix ont été motivés par la nécessité d'analyser des images acquises en imagerie par résonance magnétique sur des patients atteints de sclérose en plaques. Cette maladie est encore très méconnue tant dans sa genèse que sur les causes des handicaps qu'elle peut induire. De plus aucun traitement efficace n'est connu à l'heure actuelle. Le besoin de valider des hypothèses sur les lésions de sclérose en plaque nous a conduit à proposer des méthodes de groupement de séries temporelles ne nécessitant pas d'a priori sur le résultat final, méthodes encore peu développées en traitement d'images.

Abstract

Due to the dramatic increase of longitudinal acquisitions in the past decades such as video sequences, global positioning system (GPS) tracking or medical follow-up, many applications for time-series data mining have been developed. Thus, unsupervised time-series data mining has become highly relevant with the aim to automatically detect and identify similar temporal patterns between time-series. In this work, we propose a new spatio-temporal filtering scheme based on the mean-shift procedure, a state of the art approach in the field of image processing, which clusters multivariate spatio-temporal data. We also propose a hierarchical time-series clustering algorithm based on the dynamic time warping measure that identifies similar but asynchronous temporal patterns. Our choices have been motivated by the need to analyse magnetic resonance images acquired on people affected by multiple sclerosis. The genetics and environmental factors triggering and governing the disease evolution, as well as the occurrence and evolution of individual lesions, are still mostly unknown and under intense investigation. Therefore, there is a strong need to develop new methods allowing automatic extraction and quantification of lesion characteristics. This has motivated our work on time-series clustering methods, which are not widely used in image processing yet and allow to process image sequences without prior knowledge on the final results.

Remerciements

Je remercie tout d'abord Hugues BENOIT-CATTIN, Professeur à l'INSA Lyon, pour avoir dirigé ce travail et m'avoir fait profiter de ses précieux conseils et de son expérience tout au long de mon doctorat.

Je tiens à remercier tout particulièrement Thomas GRENIER, Maître de Conférences à l'INSA Lyon, pour m'avoir encadré pendant ce travail. Je le remercie pour sa disponibilité, son écoute, ses nombreux conseils ; pour m'avoir transmis son goût de la rigueur scientifique et accordé sa confiance tout au long de ces trois années. Nous avons partagé plus qu'une simple relation de travail : les parties de tennis, de foot, de pêche, le terroir...

Je remercie aussi Christophe DUCOTTET, Professeur à Télécom Saint Etienne, qui a accepté de présider le jury de thèse. Je le remercie pour les enseignements dispensés lors de mon cursus à Télécom Saint Etienne. Il a ainsi contribué à développer mon intérêt et mes compétences en traitement d'images et, par extension, à la réussite de ce doctorat.

Je remercie très sincèrement Ahlame DOUZAL, Maître de Conférences HDR à l'université Joseph Fourier, et Ludovic MACAIRE, Professeur à l'université de Lille, pour avoir accepté d'être les rapporteurs de ce travail. Je les remercie pour le temps qu'ils ont consacré à la lecture du manuscrit, leurs remarques avisées et la qualité de nos échanges, qui ont permis d'en faire ce qu'il est.

Merci à tous mes collègues de De Vinci, Mickaël, Fred, Sorina, Carole, David, Guy et les autres pour les échanges scientifiques, les repas de Noël, les friday cake... Merci aussi à tous mes co-llègues-doctorants-stagiaires pour les bons moments passés à et en dehors de CREATIS, Eric, Hugo, Mathilde, Pierre-Antoine, Yue, Ricardo, Arnaud, Laure, Emmanuelle, Anaïs, Diandra. Sans oublier les membres de l'équipe images et modèles qui m'ont accueilli et tous les autres (enseignants)-chercheurs que j'ai pu côtoyer durant ces trois années.

Un grand merci à ma famille, à mes parents qui m'ont toujours soutenu et été fiers de moi même dans les moments de doute.

Et enfin une attention toute particulière pour Louise, parce qu'elle est là, qu'elle m'a supporté ces trois dernières années, avec des horaires et des humeurs aléatoires... et avec qui le plus important reste à écrire.

Table des matières

Liste des figures

Liste des tableaux

Liste des algorithmes

I Introduction générale

Ce travail de thèse a été motivé par des questions posées autour d'une pathologie particulière, la sclérose en plaques (cf. chapitre 12.3) et le rôle de l'imagerie IRM pour la compréhension des mécanismes physiopathologiques sous-jacents, comme pour le suivi et la prise en charge des patients. Le but est d'identifier comment suivre, caractériser, comprendre des phénomènes variables dans l'espace et dans le temps à partir de données issues d'images. Ces problématiques nous ont amené à nous positionner sur la classification non supervisée de données spatio-temporelles multidimensionnelles et de ses applications en imagerie. L'objectif de ces méthodes de fouille de données est d'identifier des groupes d'observations similaires ; *i.e.* en traitement d'images grouper les pixels (en 2D) ou les voxels (en 3D) qui se ressemblent par rapport à une mesure de (dis)similarité donnée. L'objectif revient à maximiser la similarité des observations appartenant à un même groupe tout en maximisant la dissimilarité entre différents groupes, sans a priori sur le nombre ou la constitution des groupes à retrouver. Nous avons donc choisi de nous positionner sur des méthodes de *classification non supervisée/clustering/groupement* de données spatio-temporelles, car la sclérose en plaques est encore très méconnue des spécialistes de la santé. Sa génèse, l'origine des symptômes dont souffrent les patients ainsi que les traitements qui permettraient de la soigner sont des sujets qui suscitent beaucoup d'interrogations. Nous sommes donc dans le cas d'études prospectives de phénomènes encore méconnus, par conséquent l'introduction d'a priori dans les méthodes d'analyse n'est pas approprié car ils pourraient biaiser les conclusions obtenues.

La deuxième partie de ce manuscrit propose une étude des mesures et méthodes existantes dans le contexte général du groupement de séries temporelles. Nous étudions au premier chapitre les différents types de mesures utilisées pour comparer des séries temporelles. Le chapitre suivant est consacré à l'analyse des méthodes classiquement utilisées pour le groupement des séries temporelles. Ces études permettent de justifier les orientations méthodologiques, notamment le choix de l'approche *mean-shift* qui répond aux enjeux de faible a priori et dont l'extension aux séries temporelles est abordée dans la partie suivante.

La troisième partie de ce manuscrit détaille nos contributions méthodologiques pour le groupement des séries temporelles appliqué au traitement d'images. Nous étendons le formalisme de la méthode *mean-shift* pour qu'elle puisse prendre en compte des données temporelles. Cette méthode basée sur l'estimation de densité dans l'espace des caractéristiques est néanmoins sensible aux modulations des signaux sur l'axe temporel. C'est pourquoi nous proposons ensuite un algorithme de groupement utilisant la mesure de *dynamic time warping* (DTW) pour surmonter cette limitation. N'ayant pas défini de manière explicite l'utilisation des valeurs temporelles dans les formalismes précédents, nous basant uniquement sur le caractère séquentiel des données, nous proposons une formulation de *mean-shift* qui utilise explicitement les valeurs temporelles des acquisitions.

La quatrième et dernière partie détaille les résultats applicatifs des méthodes que nous avons proposées. Plus précisément, nous les utilisons pour l'analyse d'images acquises par résonance magnétique sur des patients atteints de sclérose en plaques. Nous commençons

par un rappel des principes physiques utilisés en imagerie par résonance magnétique (IRM). Nous abordons ensuite plus précisément la sclérose en plaques (SEP), ses causes, l'avancement de la recherche sur ce sujet et les enjeux de nos contributions dans ce domaine.

Nous concluons ce manuscrit en rappelant nos contributions méthodologiques, en soulignant les résultats applicatifs obtenus et en dégageant les perspectives suite à nos travaux.

II Etat de l'art : groupement de séries temporelles

Chapitre 1

Introduction

Le clustering, ou groupement, de séries temporelles nécessite de faire face à deux grandes problématiques. Dans un premier temps, il faut choisir une mesure de comparaison des séquences à analyser en fonction des invariances nécessaires par rapport à l'application : facteur d'échelle en amplitude ou en temps, décalage en amplitude, déphasage temporel global ou distorsions temporelles locales, occultations de mesures ou complexité des séries temporelles [1]. Dans un second temps, il faut choisir un algorithme de groupement des séries temporelles afin de retrouver les différents ensembles constituant les données étudiées.

Dans le chapitre 2, nous abordons les différentes familles de mesures de comparaison entre séries temporelles. Nous mettons en avant leurs avantages et inconvénients vis à vis de leur complexité algorithmique ou des invariances leur étant associées.

Nous étudions ensuite les méthodes classiques utilisées pour le groupement de séries temporelles au chapitre 3. Leurs facilités de paramétrage ainsi que les hypothèses induites par l'utilisation de ces méthodes sur les données traitées sont aussi exposées. Nous concluons sur leurs intérêts respectifs pour le groupement de séries temporelles extraites d'images réelles acquises au cours du temps et mettons en perspective les choix ayant motivé l'utilisation de la méthode mean-shift présentée au chapitre 4. Nos contributions méthodologiques basées sur cette méthode sont ensuite présentées en partie III.

Chapitre 2

Mesures de comparaison des séries temporelles

2.1 Introduction

Dans ce chapitre, nous présentons les grandes familles de mesures utilisées pour la comparaison des séries temporelles. Nous abordons ces mesures en étudiant leurs propriétés d'invariances, leurs limites ainsi que leurs complexités algorithmiques en temps. Nous commençons par exposer en section 2.2 les notations et définitions nécessaires à la compréhension du chapitre. La première famille présentée est celle des normes L_p (section 2.3), suivra ensuite les mesures basées sur la compression (section 2.4) puis les mesures basées sur la distance de Levenshtein (section 2.5). Nous terminons ce chapitre en section 2.6, avec les mesures basées sur la corrélation temporelle entre les séries.

2.2 Notations et définitions

Nous rappelons ici les notations et définitions nécessaires à la compréhension des mesures présentées dans ce chapitre. Considérons deux séries temporelles, ou séquences, $\mathbf{u}$ et $\mathbf{v}$. Ces vecteurs contiennent des éléments u_t et v_t, avec $t = 1 \dots \mathrm{T}$, représentant des mesures mono ou multidimensionnelles ordonnées temporellement et de dimension T. Rappelons qu'il est appelé *mesure de similarité* une fonction dont la valeur augmente avec la similarité des données. A contrario, est appelée *mesure de dissimilarité* une mesure dont la valeur augmente avec la dissimilarité des données. Ceci dit, une *distance* est une mesure de dissimilarité satisfaisant les conditions suivantes :

TABLE 2.1 – Propriétés à satisfaire pour qu'une application d définie sur $\mathrm{T}\times\mathrm{T}$ soit considérée comme une distance.

Propriété	**Définition mathématique**
Non négativité	$d : \mathrm{T}\times\mathrm{T} \rightarrow \mathbb{R}^+$
Symétrie	$\forall(\mathbf{u},\mathbf{v}) \in \mathrm{T}^2,\ d(\mathbf{u},\mathbf{v}) = d(\mathbf{v},\mathbf{u})$
Séparation	$\forall(\mathbf{u},\mathbf{v}) \in \mathrm{T}^2,\ d(\mathbf{u},\mathbf{v}) = 0 \Leftrightarrow \mathbf{u} = \mathbf{v}$
Inégalité triangulaire	$\forall(\mathbf{u},\mathbf{v},\mathbf{w}) \in \mathrm{T}^3,\ d(\mathbf{u},\mathbf{w}) \leq d(\mathbf{u},\mathbf{v}) + d(\mathbf{v},\mathbf{w})$

La fonction $d(\mathbf{u},\mathbf{v})$ doit par conséquent être non négative, séparable, symétrique et satisfaire l'inégalité triangulaire afin d'être considérée comme une distance. Un cas particulier relaxant la contrainte de séparation à $d(\mathbf{u},\mathbf{u}) = 0$ permet de définir les mesures appelées *pseudo métriques* tandis que celles ne satisfaisant pas la contrainte d'inégalité triangulaire sont appelées *semi métriques*.

2.3 La distance euclidienne

La distance euclidienne est admise comme la manière la plus simple de comparer deux séries temporelles. Cette mesure est bien une distance car elle satisfait les conditions énoncées tableau 2.1. Elle est définie comme ceci :

$$L_2(\mathbf{u},\mathbf{v}) = \sqrt{\sum_{t=1}^{\mathrm{T}} |u_t - v_t|^2} \tag{2.1}$$

Bien que l'ordonnancement temporel des mesures n'importe pas, *i.e.* l'ordre des acquisitions n'affecte en rien la mesure tant qu'il est identique entre les deux séries. Elles doivent néanmoins contenir le même nombre d'éléments T.

La complexité de cette distance est de l'ordre de $\Theta(\mathrm{T})$. Plus précisément, sa complexité est de l'ordre de $\Theta(min(\mathrm{T}_u, \mathrm{T}_v))$, T_u et T_v étant respectivement les nombres de mesures acquises pour les séries $\mathbf{u}$ et $\mathbf{v}$. Cela implique la suppression de $|\mathrm{T}_u - \mathrm{T}_v|$ mesures de la plus grande série temporelle, car cette métrique impose que les deux séries temporelles soient de tailles identiques. Par conséquent, le fait de ne pas exploiter l'ensemble des données acquises ne la rend pas optimale dans ce cas de figure.

De plus, la pertinence de la distance euclidienne s'arrête lorsque les évolutions de mesures ne sont pas synchrones. Un léger décalage entre deux séries peut entrainer une hausse de leur dissimilarité alors qu'il ne s'agit que d'un décalage d'évolution des mesures, comme présenté dans l'exemple suivant :

$$\begin{aligned} \mathbf{u} &= [1\,1\,2\,5\,3\,2\,1\,1\,0\,0] \\ \mathbf{v} &= [0\,0\,1\,1\,2\,5\,3\,2\,1\,1] \end{aligned} \qquad \Rightarrow \qquad L_2(\mathbf{u},\mathbf{v}) = 6 \tag{2.2}$$

La distance euclidienne n'est donc pas appropriée quand l'évolution des mesures subit des distorsions temporelles, ici un simple décalage temporel dans l'évolution des mesures.

Lorsque la taille des séries temporelles augmente, leur similarité peut augmenter. Néanmoins, cela ne va pas vers une diminution de la distance euclidienne :

$$\begin{aligned}\mathbf{u} &= [1\,1\,2\,5\,3\,2\,1\,1\,0\,0\,0\,0\,0\,0\,1\,1\,2\,5\,3\,2\,1\,1\,0\,0]\\ \mathbf{v} &= [0\,0\,1\,1\,2\,5\,3\,2\,1\,1\,0\,0\,0\,0\,1\,1\,2\,5\,3\,2\,1\,1\,0\,0]\end{aligned} \qquad \Rightarrow \qquad L_2(\mathbf{u},\mathbf{v}) = 6 \tag{2.3}$$

On voit dans les exemples 2.2 et 2.3 que la distance euclidienne ne prend pas du tout en compte la taille de $\mathbf{u}$ et $\mathbf{v}$. Or, dans le deuxième cas on pourrait s'attendre à une baisse de la mesure de dissimilarité car les valeurs ajoutées aux séries temporelles sont identiques. Par conséquent, on peut se questionner sur la pertinence de la distance euclidienne, sensible à de faibles distorsions et ignorant la taille des séries, bien qu'elle reste très utilisée en groupement de séries temporelles. Pour terminer, la distance euclidienne n'est qu'un cas particulier des normes L_p, applicables de la même manière aux séries temporelles, où $p = 2$:

$$L_p(\mathbf{u},\mathbf{v}) = \sqrt[p]{\sum_{t=1}^{\mathrm{T}} |u_t - v_t|^p} \tag{2.4}$$

Bien que la distance euclidienne, *i.e.* la norme L_2, soit la plus populaire pour la comparaison de séries temporelles les normes L_1 et $L_\infty(\mathbf{u},\mathbf{v}) = \max_{t=1}^{\mathrm{T}}(|u_t - v_t|)$ ont aussi déjà été utilisées dans la littérature. Yi *et al.* [2] par exemple, illustrent l'impact de leur choix sur le plus proche voisin retrouvé pour une séquence avec chacune de ces mesures.

2.4 Mesures basées sur la compression

Les mesures basées sur la compression sont issues d'un article fondateur de Kolmogorov [3]. Utilisées pour comparer des données, elles reviennent à chercher le plus petit programme capable de les générer. Cette théorie est appelée "complexité de Kolmogorov" [4] qui dans notre cas est approximée par l'utilisation d'un algorithme de compression, le dictionnaire construit par l'algorithme de compression étant le programme dans la théorie de Kolmogorov.

Une mesure de dissimilarité entre séries temporelles basée sur la compression a été proposée par Keogh *et al.* [5] :

$$CDM(\mathbf{u},\mathbf{v}) = \frac{C(\mathbf{uv})}{C(\mathbf{u}) + C(\mathbf{v})} \tag{2.5}$$

où $C(\mathbf{u})$, $C(\mathbf{v})$ sont respectivement les tailles des séquences $\mathbf{u}$ et $\mathbf{v}$ après compression et où $C(\mathbf{uv})$ est la taille compressée de la concaténation de ces deux séquences. Il est évident que la qualité de cette mesure dépend de l'algorithme de compression utilisé. Keogh *et al.* [5] ont fait le choix d'utiliser l'algorithme qui offre la meilleur compression des données, mais il reste important de bien choisir le mode de représentation des données car celui-ci impacte directement la qualité du résultat. Ceci étant, ce type de mesure sera d'autant plus approprié que les séquences sont longues, autrement leur intérêt est limité

dans la mesure où les algorithmes de compression approximent moins bien la complexité de Kolmogorov. Pour finir, ces mesures se comportent comme des boîtes noires donnant une information sur la dissimilarité globale entre deux séquences sans donner d'information sur les similarités ayant permis de compresser les données. Il faudrait envisager une étude des tables ou des dictionnaires générés pour la compression des séquences, permettant par exemple d'identifier a posteriori les ensembles ou suites d'éléments utilisés pour le calcul de la mesure. C'est d'ailleurs sur ce point que sont basées les mesures présentées dans la section suivante, qui vont chercher à déterminer les sous-ensembles communs d'éléments pour calculer une (dis)similarité.

2.5 Plus longue sous-séquence commune et dynamic time warping

Les mesures de plus longue sous-séquence commune (PLSC) et de dynamic time warping (DTW) sont toutes deux issues de la distance de Levenshtein [6], ou plus communément appelée distance d'édition. Cette dernière formalise la notion de distance entre deux chaines de caractères comme le nombre minimal d'insertions, de suppressions ou de changements de caractères nécessaires pour les rendre identiques. Bien que la distance d'édition satisfasse les contraintes décrites Table 2.1 ce n'est pas le cas des mesures PLSC et DTW, les raisons seront détaillées dans les paragraphes suivants. Néanmoins, comme la distance d'édition, elles prennent toutes deux en compte l'ordonnancement temporel des séquences pour déterminer la valeur de la mesure et sont robustes aux distorsions sur l'axe temporel pour le calcul de (dis)similarité.

2.5.1 Plus longue sous-séquence commune

La PLSC est une variation de la distance d'édition visant à identifier la plus grande suite de caractères commune à deux chaines de caractères. Cette mesure a été adaptée par Vlachos *et al.* [7] afin d'identifier la plus longue sous-séquence commune entre deux séquences numériques. Il s'agit d'une mesure de similarité car sa valeur augmente proportionnellement à la taille de la sous-séquence commune.

Le calcul de la PLSC est un problème récursif qui, abordé en *force brute* est de complexité exponentielle en temps. Il nécessite, pour chaque sous-séquence, de déterminer la valeur de sa plus grande sous-séquence commune et ainsi de suite pour chaque sous séquence. Ce problème peut être résolu efficacement grâce à la programmation dynamique en commençant par le plus petit sous problème $PLSC(u_1, v_1)$ et en utilisant son résultat pour déterminer les tailles des plus longues sous-séquences communes de tailles supérieures, *i.e.* $PLSC(\mathbf{u}_{1,2}, v_1) \dots PLSC(\mathbf{u}_{1,\dots,\mathrm{T}_u}, \mathbf{v}_{1,\dots,\mathrm{T}_v})$. La formulation suivante permet de déterminer la taille de la PLSC entre deux séries temporelles $\mathbf{u}$ et $\mathbf{v}$ pour tous les couples de sous

séquences possibles, via une matrice de coûts C de taille $\mathrm{T}_u \times \mathrm{T}_v$:

$$C(i,j) = \begin{cases} 0 & \text{si } i=0 \text{ ou } j=0 \\ C[i-1,j-1]+1 & \text{si } i,j>0 \text{ et } |i-j| \leq \delta \text{ et } u_i - v_j \leq \epsilon \\ max(C[i,j-1], C[i-1,j]) & \text{si } i,j>0 \text{ et } |i-j| \leq \delta \text{ et } u_i - v_j > \epsilon \end{cases} \quad (2.6)$$

avec ϵ l'écart maximal pour que deux mesures soit considérées comme égales et $\delta \in [0; max(T_u, T_v)]$ le paramètre contrôlant la flexibilité de recherche dans le domaine temporel.

La complexité en temps du calcul de la matrice de coûts de la PLSC est maintenant de l'ordre $\Theta(\delta.max(\mathrm{T}_u, \mathrm{T}_v))$ ou $\Theta(\delta \mathrm{T})$ si $\mathrm{T}_u = \mathrm{T}_v$ (Petitjean [8]). On peut noter que le calcul de la PLSC ne nécessite pas forcément deux séries temporelles de tailles identiques, *i.e.* T_u peut être différent de T_v. Une illustration de son calcul entre deux séries $\mathbf{u}$ et $\mathbf{v}$ est présentée Fig. 2.1. Si aucune contrainte sur le fenêtrage temporel δ est utilisée, sa complexité est de l'ordre $\Theta(\mathrm{T}_u \mathrm{T}_v) \Rightarrow \Theta(\mathrm{T}^2)$.

La taille de la plus longue sous-séquence commune est égale au nombre stocké dans le dernier indice calculé de la matrice de coûts $C(\mathrm{T}_u, \mathrm{T}_v)$. Il est facile de retrouver quelle est la plus longue sous-séquence commune entre $\mathbf{u}$ et $\mathbf{v}$. Il suffit de rebrousser le chemin partant de $C(\mathrm{T}_u, \mathrm{T}_v)$, si pour chaque case de la matrice de coûts l'élément ayant permis d'obtenir l'élément courant a été mémorisé. Si ce n'est pas le cas, il existe trois possibilités pour déterminer qui de $C[i-1,j-1], C[i,j-1]$ ou $C[i-1,j]$ a permis de calculer $C[i,j]$. Si jamais $C[i,j-1] = C[i-1,j]$, le choix de l'élément retenu pour le calcul et la construction de la PLSC dépend de la manière dont est programmée la mesure. Par conséquent, la PLSC retourne la valeur d'une des plus longues sous-séquences communes en fonction de son implémentation. La complexité en temps pour retrouver les éléments de la PLSC est linéaire et égale à $\Theta(\mathrm{T}_u + \mathrm{T}_v)$ (Cormen *et al.* [9]). Nous appellerons l'ensemble des couples $C(i,j)$ permettant de reconstruire la PLSC $P^* = \{p_l^*\}_{l=1...L}$ avec $p_1^* = C(\mathrm{T}_u, \mathrm{T}_v)$.

Vlachos *et al.* [7] ont aussi proposé une mesure de dissimilarité normalisée temporellement, basée sur la PLSC :

$$Dissim_{\delta,\epsilon}(\mathbf{u},\mathbf{v}) = 1 - \frac{PLSC_{\delta,\epsilon}(\mathbf{u},\mathbf{v})}{min(\mathrm{T}_u, \mathrm{T}_v)} \quad (2.7)$$

Le fait que $D_{\delta,\epsilon}(\mathbf{u},\mathbf{v})$ soit une mesure de dissimilarité n'en fait cependant pas une distance. Prenons par exemple :

$$\begin{cases} \mathbf{u} & = [1\ 2] \\ \mathbf{v} & = [1\ 2\ 3] \\ \mathbf{w} & = 3 \end{cases} \Rightarrow \begin{cases} PLSC_{\delta,\epsilon}(\mathbf{u},\mathbf{w}) & = 0 \\ PLSC_{\delta,\epsilon}(\mathbf{u},\mathbf{v}) & = 2 \\ PLSC_{\delta,\epsilon}(\mathbf{w},\mathbf{v}) & = 1 \\ \delta = 0.5, \epsilon = 3 & \end{cases} \Rightarrow \begin{cases} Dissim_{\delta,\epsilon}(\mathbf{u},\mathbf{w}) & = 1 \\ Dissim_{\delta,\epsilon}(\mathbf{u},\mathbf{v}) & = 0 \\ Dissim_{\delta,\epsilon}(\mathbf{w},\mathbf{v}) & = 0 \\ \delta = 0.5, \epsilon = 3 & \end{cases} \quad (2.8)$$

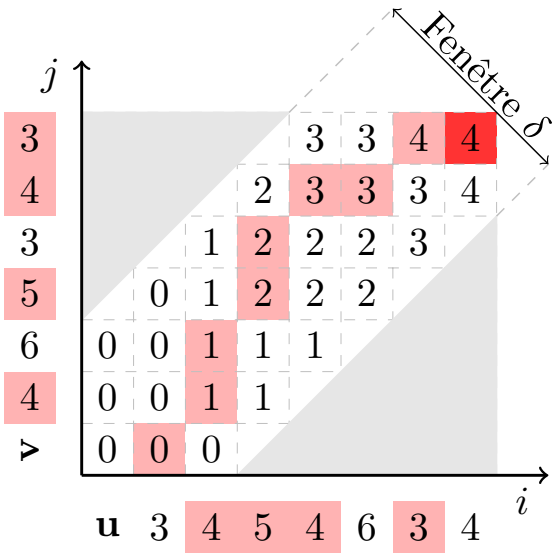


FIGURE 2.1 – Exemple de matrice de coûts pour le calcul de la PLSC entre deux séquences. Paramétrage de $\delta = 2 \Leftrightarrow |i-j| \leq 2$ et $\epsilon = 0.5$. En rouge : chemin P^* reconstruit pour retrouver une des plus longues sous-sequences communes entre $\mathbf{u}$ et $\mathbf{v}$. La valeur de la PLSC apparaît en rouge foncé en position p_1^*. La PLSC, mise en évidence sur les séries $\mathbf{u}$ et $\mathbf{v}$, apparaissant à côté des axes des abscisses et des ordonnées est 4 5 4 3 et sa valeur est égale à 4.

$D_{\delta,\epsilon}(\mathbf{u}, \mathbf{v})$ est une *semi pseudo-métrique* dans la mesure où c'est une *pseudo-métrique* (cf. section 2.2) qui ne satisfait pas la contrainte de l'inégalité triangulaire.

2.5.2 Dynamic time warping

La DTW est une mesure de dissimilarité, plus particulièrement une *semi pseudo-métrique*, introduite pas Sakoe et Chiba [10, 11] dont le but original était de comparer et de reconnaitre des signaux vocaux. La démonstration de cette qualification est basée sur un exemple du même type que (2.8). Cette mesure est reconnue comme un standard pour la classification et le groupement de séries temporelles [12, 13].

Le calcul de la DTW permet d'apparier/d'aligner de manière optimale les éléments de deux séries temporelles. Il est important de noter que les couples de valeurs appariées par la DTW respectent l'ordonnancement temporel des éléments, en d'autres termes les appariements ne peuvent pas "se croiser" temporellement. En plus de cette contrainte dite de monotonicité, les alignements doivent aussi être continus ce qui signifie, contrairement à la PLSC, qu'aucun élément des séries temporelles ne peut être ignoré. Si nous considérons $P^*_{\mathbf{u},\mathbf{v}} = \{\mathbf{p}_l^*\}_{l=1...\mathrm{L}}$ comme les couples d'indices $\{(i,j) \mid 1 \leq i \leq \mathrm{T}_u, 1 \leq j \leq \mathrm{T}_v\}$ représentant l'ensemble des appariements optimaux/le chemin d'alignements entre deux séries temporelles $\mathbf{u}$ et $\mathbf{v}$, ces deux contraintes se traduisent par $(0,0) < \mathbf{p}_l^* - \mathbf{p}_{l-1}^* \leq (1,1)$. La détermination d'un des chemins d'alignements optimal $P^*_{\mathbf{u},\mathbf{v}}$ résulte de la minimisation

suivante :

$$P^*_{\mathbf{u},\mathbf{v}} = arg\min_{P} \left(\frac{\sum_{l=1}^{\mathrm{L}} d(\mathbf{u},\mathbf{v},\mathbf{p}_l)\cdot w_l}{\sum_{l=1}^{\mathrm{L}} w_l} \right) \tag{2.9}$$

avec w_l les poids associés à chaque appariement et d la distance euclidienne entre deux éléments associés :

$$d(\mathbf{u},\mathbf{v},\mathbf{p}_l) = \| \mathbf{p}_l(\mathbf{u}) - \mathbf{p}_l(\mathbf{v}) \| \tag{2.10}$$

Nous pouvons déduire à partir de P^* les versions alignées de $\mathbf{u}$ et $\mathbf{v}$ notées respectivement $P^*_{\mathbf{u},\mathbf{v}}(\mathbf{u})$ et $P^*_{\mathbf{u},\mathbf{v}}(\mathbf{v})$. Simplifions les notations de ces deux vecteurs en $P^*(\mathbf{u})$ et $P^*(\mathbf{v})$ pour plus de légèreté. P^* est composé de L couples appariés, $\mathbf{p}^*_l(\mathbf{u})$ et $\mathbf{p}^*_l(\mathbf{v})$ correspondant aux $l^{ièmes}$ mesures des séquences $\mathbf{u}$ et $\mathbf{v}$ une fois alignées. Par conséquent, le formalisme de la DTW est le suivant :

$$DTW(\mathbf{u},\mathbf{v}) = \min_{\mathbf{p}_l} \left(\frac{\sum_{l=1}^{\mathrm{L}} d(\mathbf{u},\mathbf{v},\mathbf{p}_l)\cdot w_l}{\sum_{l=1}^{\mathrm{L}} w_l} \right) = \frac{\sum_{l=1}^{\mathrm{L}} d(\mathbf{u},\mathbf{v},\mathbf{p}^*_l)\cdot w_l}{\sum_{l=1}^{\mathrm{L}} w_l} \tag{2.11}$$

A l'instar de la PLSC, ce problème de minimisation peut se résoudre de manière efficace grâce à la programmation dynamique via le calcul d'une matrice de coûts C de taille $\mathrm{T}_u \times \mathrm{T}_v$ dont les éléments sont calculés de la sorte :

$$C(i,j) = \begin{cases} 2\,\|\,u_1 - v_1\,\| & \text{si } i=1 \text{ et } j=1 \\ \min \begin{cases} C[i,j-1] + \|\,u_i - v_i\,\| \\ C[i-1,j] + \|\,u_i - v_i\,\| \\ C[i-1,j-1] + 2\,\|\,u_i - v_i\,\| \end{cases} & \text{si } i \text{ ou } j > 1 \text{ et } |i-j| \leq \delta \end{cases} \tag{2.12}$$

avec δ le paramètre de fenêtrage temporel, similaire à celui introduit dans le calcul de la PLSC (2.6). Dans ces équations, les valeurs multipliant $\|\,u_i - v_i\,\|$ correspondent aux poids w_l (1, 1 et 2) apparaissant en (2.9) et (2.11). La valeur de la DTW, comme proposée par Sakoe et Chiba [11], n'est pas directement égale à $C(\mathrm{T}_u, \mathrm{T}_v)$ mais est donnée par sa valeur normalisée par la somme des pondérations w_l :

$$DTW(\mathbf{u},\mathbf{v}) = \frac{C(\mathrm{T}_u, \mathrm{T}_v)}{\sum_l w_l} \tag{2.13}$$

Une fois la matrice de coûts calculée, il est simple de reconstruire un chemin d'alignements en partant de l'indice $(\mathrm{T}_u, \mathrm{T}_v)$ de la matrice C, comme pour la PLSC. Il faut noter les deux contraintes nécessaires à la reconstruction du chemin d'alignements : $\mathbf{p}^*_1 = (\mathrm{T}_u, \mathrm{T}_v)$ et $\mathbf{p}^*_{\mathrm{L}} = (1,1)$. Un exemple de calcul de la DTW entre deux séries temporelles $\mathbf{u}$ et $\mathbf{v}$ ainsi que les appariements optimaux trouvés après reconstruction du chemin d'alignements sont

présentés Fig. 2.2. Dans le cas particulier où $\delta = 0$, le calcul de la DTW revient au calcul de la distance euclidienne. Le seul chemin d'alignements autorisé revient à calculer les différences temps à temps, $i = j$, entre les séries temporelles.

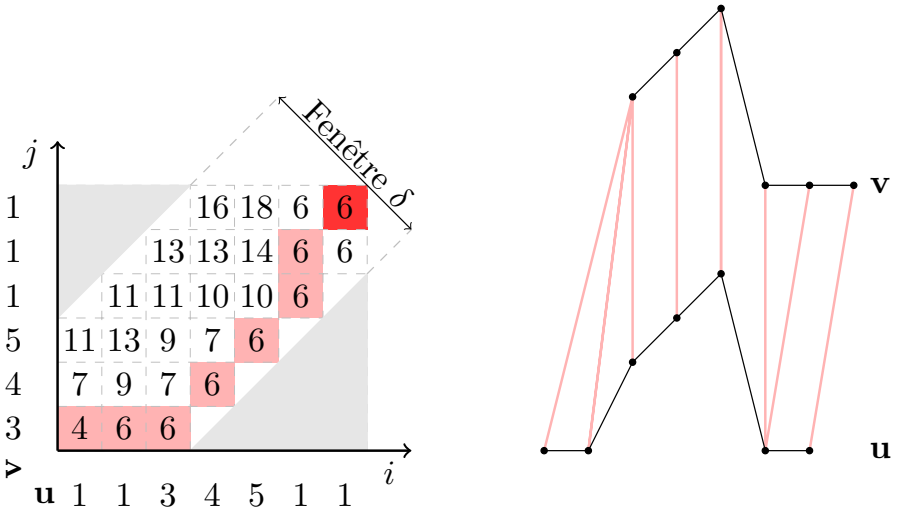


FIGURE 2.2 – A gauche : matrice de coûts de la DTW avec $\delta = 2 \Leftrightarrow |i-j| \leq 2$. A droite : associations trouvées par la DTW après minimisation du coût de trajet entre les éléments $\mathbf{p}_1^* = (\mathrm{T}_u, \mathrm{T}_v)$ et $\mathbf{p}_\mathrm{L}^* = (1,1)$ de la matrice C (2.9). L'absence d'axes est voulu ainsi que le décalage horizontal et vertical des courbes afin de faciliter la visibilité des valeurs appariées. Néanmoins, les dynamiques des amplitudes des séquences **u** et **v** sont à l'échelle. Le chemin d'alignements est identifié en rouge dans la matrice de coûts et les alignements optimaux retrouvés apparaissent en rouge sur le schéma de droite. Il s'agit d'un des chemins d'alignements optimal, le choix est fait en fonction de l'implémentation de la DTW.

Brièvement, l'analyse de la complexité en temps de la DTW est identique à celle présentée pour la PLSC. Elle est de l'ordre de $\Theta(\delta.max(\mathrm{T}_u, \mathrm{T}_v))$ ou $\Theta(\delta \mathrm{T})$ si $\mathrm{T}_u = \mathrm{T}_v$ (Petitjean [8]) pour le calcul de la valeur de la DTW et de l'ordre $\Theta(\mathrm{T}_u + \mathrm{T}_v)$ pour la reconstruction du chemin d'alignements (Cormen *et al.* [9]). D'autres formes géométriques contraignant la zone de recherche du chemin d'alignements optimal ont été proposées, réduisant toujours la complexité calculatoire de la DTW, comme par exemple le parallélogramme d'Itakura [14]. Les contraintes sur les zones de recherche, *i.e.* sur les appariements autorisés et leurs importances dans le calcul de la DTW, sont définis via des valeurs de poids w_l spécifiques à l'approche choisie.

La mesure DTW, comme la PLSC, ne contient pas de normalisation en amplitude des séries temporelles dans son formalisme permettant d'être invariant aux facteurs d'échelle en amplitude. Par conséquent, bien qu'insensibles aux déphasages temporels, ces mesures ne seront pas robustes en cas de trop fortes dissimilarités entre les dynamiques en amplitude des séries temporelles. Cette limitation peut être surmontée par une pré-normalisation des données temporelles ou bien par l'utilisation de métriques incluant une normalisation dans leurs formalismes, comme celles présentées dans la section suivante.

2.6 Dissimilarité en comportement et valeur

Cette mesure de dissimilarité entre séries temporelles a été introduite par Douzal Chouakria *et al.* [15] et prend en compte deux informations différentes : (1) la similarité des comportements (variations des mesures similaires et synchronisées, opposées ou indépendantes) ; (2) l'écart existant entre les mesures/valeurs des série temporelle.

La similarité des comportements entre deux séries temporelles $\mathbf{u}$ et $\mathbf{v}$ est mesurée via le calcul d'une extension de la corrélation de Pearson aux mesures temporelles :

$$Cort(\mathbf{u},\mathbf{v}) = \frac{\sum_{t=1}^{\mathrm{T}-1}(u_{t+1}-u_t)(v_{t+1}-v_t)}{\sqrt{\sum_{t=1}^{\mathrm{T}-1}(u_{t+1}-u_t)^2}\sqrt{\sum_{t=1}^{\mathrm{T}-1}(v_{t+1}-v_t)^2}} \tag{2.14}$$

avec T la taille des séries temporelles. Cette définition du coefficient de corrélation temporelle est limitée aux différences entre éléments temporellement contigus. Une adaptation prenant en compte un chemin d'alignements quelconque P [16], comme par exemple le chemin d'alignements optimal retourné par la DTW, est définie comme :

$$Cort(\mathbf{u},\mathbf{v},P) = \frac{\sum_{l=1}^{\mathrm{L}-1}\left(\mathbf{p}_{l+1}(\mathbf{u})-\mathbf{p}_l(\mathbf{u})\right)\left(\mathbf{p}_{l+1}(\mathbf{v})-\mathbf{p}_l(\mathbf{v})\right)}{\sqrt{\sum_{l=1}^{\mathrm{L}-1}\left(\mathbf{p}_{l+1}(\mathbf{u})-\mathbf{p}_l(\mathbf{u})\right)^2}\sqrt{\sum_{l=1}^{\mathrm{L}-1}\left(\mathbf{p}_{l+1}(\mathbf{v})-\mathbf{p}_l(\mathbf{v})\right)^2}} \tag{2.15}$$

Ces équations permettent d'évaluer si deux séries temporelles évoluent "dans le même sens". Par exemple, si $Cort(\cdot) = 1$ cela veut dire que $\mathbf{u}$ et $\mathbf{v}$ ont un comportement linéairement identique tandis que si $Cort(\cdot) = -1$ cela veut dire que $\mathbf{u}$ et $\mathbf{v}$ ont un comportement linéairement opposé. Le dernier cas, $Cort = 0$, signifie que les deux séries temporelles sont linéairement indépendantes. On peut noter que $Cort(\cdot) \in [-1, 1]$. Dans Douzal Chouakria *et al.* [15], la mesure de corrélation temporelle calculée en (2.14) est ensuite associée à une fonction de coût afin de pondérer son résultat par les distances entre les valeurs de chacune des séries :

$$D_k(\mathbf{u},\mathbf{v}) = \frac{2}{1+\exp\left(k\cdot Cort(\mathbf{u},\mathbf{v})\right)}\cdot L_2(\mathbf{u},\mathbf{v}) \tag{2.16}$$

avec L_2 la distance euclidienne définie en (2.1). Une extension de cette formule similaire au calcul de la DTW a été proposée [16] afin d'utiliser un chemin d'alignements optimal entre deux séries temporelles :

$$D_k(\mathbf{u},\mathbf{v}) = \min_P\left(\frac{2}{1+\exp\left(k\cdot Cort(\mathbf{u},\mathbf{v},P)\right)}\cdot L_1(P(\mathbf{u}),P(\mathbf{v}))\right) \tag{2.17}$$

$$= \frac{2}{1+\exp\left(k\cdot Cort(\mathbf{u},\mathbf{v},P^*)\right)}\cdot L_1(P^*(\mathbf{u}),P^*(\mathbf{v})) \tag{2.18}$$

considérant L_1 comme la norme L_p avec $p = 1$ définie en (2.4). Ces formulations ont l'avantage d'être robustes aux variations de dynamiques entre séries temporelles et dans le cas de la formulation (2.17) d'ajouter une robustesse à de potentiels déphasages entre séquences.

La complexité en temps de la mesure (2.16) est de l'ordre de $\Theta(T)$ tandis que celle de la mesure (2.17), en utilisant la programmation dynamique, est de l'ordre de $\Theta(T^2)$. Pour conclure sur cette mesure, on peut dire que la formulation (2.16) définit une *semi métrique* car seule la propriété d'inégalité triangulaire n'est pas remplie et la formulation (2.17) correspond à une *semi pseudo-métrique*.

D'autres mesures de dissimilarité basées sur la corrélation ont été proposées comme celle de Yang *et al.* [17], ou plus récemment celle de Paparrizos *et al.* [18]. Néanmoins, bien qu'elles soient aussi formulées pour être invariantes aux facteurs d'échelle en amplitude, ces mesures ne compensent qu'un déphasage temporel global entre les séries temporelles.

2.7 Conclusion

Dans cette partie, nous avons présenté les principales familles de mesures utilisées pour comparer les séries temporelles dans des approches de classification ou de groupement.

La distance euclidienne ou les normes de type L_p en général comparent un à un les éléments entre deux séries temporelles. Elles restent très largement utilisées en fouille de données. Les normes L_p comportent certains avantages de par leur simplicité, leur temps de calcul linéaire et le fait qu'elles ne nécessitent aucun paramétrage. Néanmoins, elles restent très sensibles au bruit et aux déphasages temporels locaux ou globaux entre les séquences. Ceci pourra s'avérer problématique pour l'étude de signaux réels. En effet, bien que certains processus évolutifs puissent être similaires, ils pourraient ne pas être considérés comme tels en cas de déphasages locaux ou globaux.

Nous avons ensuite présenté les mesures basées sur la compression. D'une part la complexité et l'efficacité de ces approches sont dépendantes de l'algorithme de compression utilisé et les résultats ne sont pas interprétables de manière simple et directe pour identifier les (dis)similarités entre les signaux temporels. Le temps n'est pas utilisé pour ordonner les données mais comme un simple attribut de chaque échantillon. Par conséquent, la non prise en compte de la séquentialité ou de sous séquences change l'ordre global des éléments, ce qui s'avère problématique. De plus, le choix du mode de représentation des données, d'un algorithme et sa paramétrisation ne sont pas triviaux et influent directement sur la qualité du résultat obtenu. Pour ces raisons, ce type de mesures n'est pas adapté à notre problématique.

Les mesures permettant de trouver des alignements élastiques entre les séries temporelles ont ensuite été abordées. Elles permettent d'imaginer que les séries temporelles puissent avoir des comportements similaires mais avec des évolutions étirées ou raccourcies. La PLSC, contrairement à la DTW, a pour particularité de pouvoir passer outre certains évènements en "sautant" des éléments dans les séries temporelles. Néanmoins, le

fait d'ignorer certaines parties des évolutions de manière non maîtrisée, car considérées comme non informatives aux yeux de la PLSC, pourrait s'avérer problématique avec un but exploratoire et de compréhension des phénomènes évolutifs étudiés. Dans cette optique la DTW pourrait se présenter comme le choix le plus judicieux : les appariements retournés alignent tous les éléments d'une série **u** à au moins un élément de la série **v** et inversement. De plus, contrairement à la PLSC, elle ne nécessite pas de déterminer un seuil pour définir la similarité entre les séries, ce qui fait économiser le réglage d'un paramètre. Ces deux mesures sont capables de trouver des alignements élastiques mais restent sensibles aux dissimilarités en amplitude entre deux séquences. Les mesures basées sur la corrélation temporelle permettent quant à elles de surmonter cette limitation.

Cependant, nous nous limiterons dans la suite des travaux à l'utilisation de la distance euclidienne ainsi que de la DTW (cf. chapitre 13) ; le problème de sensibilité aux différences de dynamiques pouvant être surmonté en normalisant les séries temporelles préalablement aux algorithmes de groupement.

C'est d'ailleurs le point que nous abordons dans la partie suivante. Après avoir étudié les principales mesures de similarité, nous présentons les méthodes classiquement utilisées pour le groupement de séries temporelles.

Chapitre 3

Méthodes de groupement des séries temporelles

3.1 Introduction

Que les données soient statiques ou dynamiques, le but des algorithmes de groupement reste le même : partitionner un certain nombre n d'observations en K groupes homogènes. Les observations, dans notre cas les séries temporelles, d'un même groupe doivent être proches tandis que celles de groupes différents doivent être éloignées au sens de la métrique utilisée. Plusieurs types d'algorithmes ont été proposés pour le groupement de séries temporelles : ceux utilisant les données temporelles brutes, ceux utilisant des caractéristiques extraites sur les données temporelles ou ceux utilisant des modèles d'évolutions. Ces derniers effectuent ensuite un groupement sur les paramètres des modèles ajustés aux séries temporelles ou sur l'écart des séries temporelles par rapport au modèle. Une taxonomie des différentes approches et une description exhaustive des méthodes utilisées dans ces différents contextes ont été proposées par Warren Liao [19].

Dans ce chapitre, nous nous intéressons plus particulièrement aux approches de partitionnement et de groupements hiérarchique et spectral qui se basent sur la forme des séries temporelles, *i.e.* qui utilisent *l'information brute* des séries temporelles pour mesurer les écarts les séparant. Jusqu'alors ces trois types d'approches étaient considérés comme les plus performants et ont été adaptés ou utilisés pour évaluer les derniers algorithmes de groupement de séries temporelles proposés (cf. Paparrizos *et al.* [18] et Begum *et al.* [13]). Par ailleurs, c'est avec la famille d'approches à laquelle appartient celle proposée par Be-

gum *et al.* [13] que nous finissons cette partie, *i.e.* les approches basées sur la densité qui présentent des propriétés intéressantes pour le groupement des séries temporelles.

3.2 Méthodes de partitionnement

Les algorithmes de partitionnement ont pour but de déterminer sans aucun ordonnancement hiérarchique des observations un nombre K de groupes, dans notre cas de séries temporelles, similaires. Le processus de partitionnement vise à optimiser une fonction de coût. Prenons un ensemble de séries temporelles $\mathbf{X} = \{\mathbf{x}_i \mid i = 1 \dots n,\, \mathbf{x}_i \in \mathbb{R}^{\mathrm{T}}\}$, les algorithmes de partitionnement ont pour put de séparer cet ensemble $\mathbf{X}$ en K groupes distincts $\mathbf{C} = \{\mathbf{C}_1 \dots \mathbf{C}_{\mathrm{K}}\}$ en maximisant ou minimisant une fonction objectif J. La partition optimisant la fonction J peut être retrouvée en *force brute* mais nécessiterait bien trop de calculs en pratique comme le montre la formule de Liu [20] :

$$P(\mathrm{N}, \mathrm{K}) = \frac{1}{\mathrm{K}!} \sum_{j=1}^{\mathrm{K}} (-1)^{(\mathrm{K}-j)} \frac{\mathrm{K}!}{j!\,(\mathrm{K}-j)!}\, j^{\mathrm{N}} \tag{3.1}$$

ce qui offrirait environ 3×10^{13} possibilités pour partitionner 30 séries en 3 groupes. Il est donc nécessaire d'utiliser des heuristiques afin d'approximer une solution.

Une méthode très répandue et classiquement utilisée est l'algorithme itératif k-means, qui vise à minimiser la somme des écarts quadratiques des observations par rapport aux centres des groupes. La fonction objectif à minimiser s'exprime donc :

$$J(\mathbf{\Gamma}, \mathbf{M}) = \sum_{j=1}^{\mathrm{K}} \sum_{i=1}^{\mathrm{n}} \gamma_{ji} L_2^2(\mathbf{x}_i, \mathbf{m}_j) \tag{3.2}$$

où

$\mathbf{\Gamma} = \{\gamma_{ji}\}$ est l'ensemble des partitions tel que $\gamma_{ji} = \begin{cases} 1 \text{ si } \mathbf{x}_i \in \mathbf{C}_j \\ 0 \text{ sinon} \end{cases}$ avec $\sum_{j=1}^{\mathrm{K}} \gamma_{ji} = 1, \forall i$;

$\mathbf{M} = \{\mathbf{m}_j\}_{j=1\dots\mathrm{K}}$ les centroïdes associés à chaque groupe $\mathbf{C}_j$; et avec

$\mathbf{m}_j = \dfrac{\sum_{i=1}^{n} \gamma_{ji}\, \mathbf{x}_i}{\sum_{i=1}^{n} \gamma_{ji}}$ les centroïdes (les barycentres) des observations associées aux groupes $\mathbf{C}_j$.

La méthode k-means, détaillée dans l'algorithme 1, a pour objectif de minimiser la somme des écarts quadratiques des observations aux centres de leurs classes respectives. Elle minimise par conséquent la variance intra-classe des partitions obtenues. Cependant, ce critère est approprié dans le cas de groupes compacts et bien séparés. Ceci peut être un problème avec des données dont la dimensionnalité augmente avec le nombre de points temporels ou l'utilisation de mesures de similarité n'étant pas des distances, donc ne satis-

Algorithm 1 Algorithme de partitionnement : k-means

Require: $\mathbf{M} = \{\mathbf{m}_j\}_{j=1...K}$: Initialisation des centroïdes aléatoirement ou avec a priori
Require: $\mathbf{\Gamma} = \{\gamma_{ji}\}$: Initialisation des partitions avec le centroïde le plus proche de chaque observation

repeat
  for all $\mathbf{x}_i \in \mathbf{X}$ **do**
    $\gamma_{oi} = arg \max_{\gamma_{ji} \in \mathbf{\Gamma}} (\gamma_{ji})$
    $\gamma_{li} = arg \min_{\mathbf{m}_j \in \mathbf{M}} (L_2(\mathbf{x}_i, \mathbf{m}_j)^2)$
    if $\gamma_{li} \neq \gamma_{oi}$ **then**
      $\gamma_{oi} = 0$
      $\gamma_{li} = 1$
    end if
  end for
  for all $\mathbf{m}_j \in \mathbf{M}$ **do**
    $\mathbf{m}_j = \sum_{i=1}^{N} \frac{\gamma_{ji}\,\mathbf{x}_i}{\gamma_{ji}}$
  end for
  Mem = $\mathbf{C}$
  Mise à jour des groupes $\mathbf{C}_j$ grâce au nouvel ensemble de partitions $\mathbf{\Gamma}$
until $C = Mem$

faisant pas la même contrainte de "sphéricité" induite par le calcul des centroïdes. De plus, ce critère est sensible aux valeurs aberrantes et au bruit ce qui peut décaler les centroïdes. Par conséquent, le partitionnement est déformé car toutes les observations doivent être associées à un groupe, même les points aberrants. L'algorithme de partitionnement autour des médoïds (PAM) a été proposé par Kaufman *et al.* [21] pour surmonter cette limitation aux valeurs aberrantes, en remplaçant les centroïdes par les observations existantes ayant la distance moyenne minimale à toutes les autres observations d'un même groupe.

La complexité de k-means est de l'ordre $\Theta(\bar{it}.\text{n.K.T})$ où $\bar{it}$ est le nombre moyen d'itérations. Comme spécifié par Xu *et al.* [22] le nombre de classes K, d'itérations $\bar{it}$ et le nombre de points temporels T sont généralement très inférieurs au nombre d'observations N ce qui induit pour k-means une complexité de l'ordre de Θ(n).

L'algorithme k-means, comme PAM, n'est pas déterministe car il converge itérativement vers un minimum local dépendant des valeurs d'initialisation des centroïdes souvent réalisé aléatoirement. De plus, il est supposé que le nombre de groupes est connu par avance ce qui n'est pas forcément le cas.

K-means a aussi été adapté avec d'autres mesures pour comparer les séries temporelles, notamment pour pouvoir retrouver des groupes dont les évolutions sont similaires mais déphasées dans le temps comme Petitjean *et al.* [23] et Soheily *et al.* [24]. Néanmoins, cela nécessite de déterminer une méthode de calcul d'une série temporelle moyenne sous conditions de distorsions temporelles pour mettre à jour les centroïdes, ce qui n'est pas trivial.

Nous allons maintenant étudier les algorithmes hiérarchiques, également utilisés pour

le groupement de séries temporelles.

3.3 Méthodes hiérarchiques

Les algorithmes hiérarchiques permettent de grouper les observations en une suite de partitions intriquées les unes aux autres. La construction de la structure hiérarchique peut se faire en mode *bottom-up* partant de singletons jusqu'à un unique groupe ou inversement (*top-down*). La construction en *bottom-up* est appelée groupement hiérarchique agglomératif tandis que la construction *top-down* est appelée groupement agglomératif divisif. Néanmoins les algorithmes hiérarchiques divisifs ne sont généralement pas utilisés car pour un ensemble de N observations, 2^{N-1} divisions possibles sont considérées au départ de la procédure ce qui engendre un nombre considérable de calculs.

Le groupement hiérarchique agglomératif peut utiliser différents critères pour déterminer les groupes les plus proches. Une étude exhaustive de ceux-ci est présentée par Xu *et al.* [22], en voici les principaux :

— Lien simple : la distance entre deux groupes est la distance la plus courte entre les membres des groupes. Un lien simple entre $\mathbf{C}_i, \mathbf{C}_j$ est exprimé comme ceci :

$$D_{SL}(\mathbf{C}_i, \mathbf{C}_j) = \min_{\mathbf{x}\in\mathbf{C}_i, \mathbf{y}\in\mathbf{C}_j} (dist(\mathbf{x}, \mathbf{y})) \tag{3.3}$$

avec *dist* une distance.

— Lien complet : la distance entre deux groupes est la distance la plus importante entre les membres des groupes. Un lien complet entre $\mathbf{C}_i, \mathbf{C}_j$ est exprimé comme ceci :

$$D_{CL}(\mathbf{C}_i, \mathbf{C}_j) = \max_{\mathbf{x}\in\mathbf{C}_i, \mathbf{y}\in\mathbf{C}_j} (dist(\mathbf{x}, \mathbf{y})) \tag{3.4}$$

— Lien moyen : la distance entre deux groupes est la distance moyenne entre toutes les paires de membres entre chaque groupe. Un lien moyen entre $\mathbf{C}_i, \mathbf{C}_j$ est exprimé comme ceci :

$$D_{AL}(\mathbf{C}_i, \mathbf{C}_j) = \frac{1}{\#\mathbf{C}_i \#\mathbf{C}_j} \sum_{i=1}^{\#\mathbf{C}_i} \sum_{j=1}^{\#\mathbf{C}_j} dist(\mathbf{x}_i, \mathbf{y}_j) \tag{3.5}$$

Sans aucune autre condition, le groupement hiérarchique s'arrête une fois qu'il n'existe plus qu'un seul groupe. Néanmoins, avec une connaissance a priori du nombre de classes K désiré, l'algorithme peut s'arrêter une fois ce critère atteint. Le problèmes majeur des algorithmes hiérarchiques est leur complexité en temps de l'ordre de $\Theta(N^2)$ ce qui peut être problématique pour des ensembles de données de tailles importantes. Ils sont aussi sensibles au bruit et aux valeurs aberrantes. Une fois qu'un groupe est formé l'algorithme ne le remettra pas en cause et si une erreur de groupement est effectuée alors elle sera propagée.

Il est intéressant de noter que les algorithmes hiérarchiques ont aussi été décrits du point de vue de la théorie des graphes (Harary [25]) dont une description détaillée a été

proposée par Jain *et al.* [26]. Ceci nous mène naturellement aux algorithme de groupement spectraux où la théorie des graphes a été utilisée pour effectuer le partitionnent des données.

3.4 Méthodes spectrales

Dans cette section, nous prenons à titre d'exemple l'algorithme proposé par Ng *et al.* [27] car il a été adapté et utilisé très récemment dans l'évaluation d'algorithmes de groupement de séries temporelles (cf. Begum *et al.* [13]). C'est aussi une approche spectrale permettant de faire un partitionnement multi-classes des observations. La suite des démarches à suivre pour grouper un ensemble d'observations $\mathbf{X} = \{\mathbf{x}_i\}_{i=1...\mathrm{N}}$ en K sous ensembles est la suivante :

1. Déterminer la matrice d'affinités $\mathbf{A} \in \mathbb{R}^{\mathrm{N}\times\mathrm{N}}$ définie comme $\mathbf{A}_{ij} = \exp(\frac{-\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2})$ avec $i \neq j$ et $\mathbf{A}_{ii} = 0$.
2. Définir $\mathbf{D} = diag\left(\sum_{j=1}^{\mathrm{N}} \mathbf{A}_{ij}\right)$ et construire la matrice laplacienne $\mathbf{L} = \mathbf{D}^{-1/2}\mathbf{A}\mathbf{D}^{-1/2}$.
3. Trouver les K premiers vecteurs propres de la matrice $\mathbf{L}$ (choisis orthogonaux en cas de valeurs propres répétées) et créer la matrice $\mathbf{V} = [\mathbf{v}_1 \dots \mathbf{v}_\mathrm{K}] \in \mathbb{R}^{\mathrm{N}\times\mathrm{K}}$, matrice des vecteurs propres concaténés dans le sens des colonnes.
4. Normaliser $\mathbf{X}$ afin d'obtenir la matrice $\mathbf{Y}$ telle que $\mathbf{Y}_{ij} = \sum_{j=1}^{\mathrm{K}} \left(\mathbf{X}_{ij}^2\right)^{1/2} \ \forall\, i \in [1;\mathrm{N}]$.
5. Considérer chaque ligne $\mathbf{Y}_{i1...\mathrm{K}} \in \mathbb{R}^\mathrm{K}$ comme une nouvelle observation dans l'espace défini par les vecteurs propres. Grouper les observations en K partitions avec k-means (ou tout autre algorithme).
6. Associer les observations initiales $\mathbf{x}_i$ au groupe k auquel a été associé le point $\mathbf{Y}_{i1...\mathrm{K}}$ lors du partitionnement.

La première étape consiste à déterminer la matrice des distances séparant les observations, pouvant être interprétée comme un graphe pondéré dont les nœuds correspondent aux observations et dont les arêtes pondérées correspondent aux distances les séparant. Dans cet algorithme, deux paramètres sont à régler : le choix de la valeur de σ permettant de pondérer ces distances pour la création de la matrice d'affinités ainsi que le nombre K de vecteurs propres et de groupes. Ng *et al.* [27] montrent que le paramétrage de σ peut être automatisé, ce qui fait un paramètre de moins a régler manuellement. Ils montrent aussi la facilité d'initialisation des centroïdes pour la phase de partitionnement. Un a priori est utilisé : le premier centroïde est choisi au hasard et les autres centroïdes devront avoir un angle proche de 90° avec tous les autres centroïdes déjà choisis, *i.e.* si les groupes sont bien séparés alors ils devraient être repartis autour d'un vecteur propre différent (par définition les vecteurs propres sont orthogonaux). Cette heuristique fonctionne très bien car les auteurs n'ont utilisé qu'une seule fois k-means pour obtenir les résultats qu'ils présentent. Néanmoins nous devons toujours faire face au choix du nombre de groupes à identifier en amont de la procédure. Nadler *et al.* [28] soulignent aussi l'incapacité de cet algorithme

à partitionner des observations formant des groupes dont les tailles ou dont les densités diffèrent. De plus, nous pouvons redouter que ce type d'approche, basée sur k-means, hérite de la difficulté des k-means à partitionner des groupes différents en présence de bruit venant "mélanger" les observations des différents groupes. Les pistes restent cependant ouvertes quant à l'utilisation d'autres algorithmes de groupement ou bien à l'utilisation d'autres métriques de similarité comme l'ont fait Begum *et al.* [13] en utilisant la DTW pour adapter cette approche au groupement de séries temporelles.

Après avoir étudié le mode de fonctionnement d'une approche classique de groupement spectral, nous allons maintenant étudier des algorithmes qui se basent non pas sur la matrice d'affinités comme définie à l'étape 1 de l'algorithme 3.4, mais sur les distances séparant les observations et la notion de densité dans l'espace des caractéristiques.

3.5 Méthodes basées sur la densité

Les algorithmes basés sur la densité ont pour but d'identifier des groupes de points sur le principe que la densité des observations appartenant à un groupe augmente en se rapprochant du centre, *i.e.* de leur représentant. Le centre du groupe étant par conséquent le point de densité maximale (ou mode). Ces algorithmes permettent aussi d'identifier des ensembles de formes non convexes, contrairement aux approches minimisant les distances entre les échantillons et les centroïdes des groupes auxquels ils sont affectés.

Ester *et al.* [29] ont proposé un algorithme de groupement appelé DBSCAN. Ils considèrent qu'une observation appartenant à un groupe doit être entourée d'un nombre minimal d'observations $MinPts$ dans un rayon ϵ. C'est la notion de densité de voisinage. Les groupes seront ainsi construits de proche en proche et les observations ne satisfaisant pas les conditions énoncées sont ignorées et considérées comme du bruit. Une adaptation de DBSCAN relâchant la contrainte ϵ, appelée OPTICS, a été proposée par Ankerst *et al.* [30]. Cependant Ertoz *et al.* [31] mettent en garde par rapport au bien fondé de l'utilisation de DBSCAN avec des séries temporelles, car selon eux la notion de densité euclidienne perd en significativité avec l'augmentation du nombre de dimensions.

Des approches inspirées de DBSCAN ont été proposées pour le groupement de trajectoires spatiales. D'après Aggarwal *et al.* [32], un des premiers travaux à avoir pris en compte l'information temporelle pour le groupement de trajectoires spatiales a été celui de Kalnis *et al.* [33] introduisant le concept de *groupe en mouvement*. Il s'agit ici d'un groupe d'observations qui se déplacent à proximité les unes des autres. C'est en quelque sorte une séquence temporelle de groupes spatiaux ayant un pourcentage minimal Θ de membres communs entre deux groupes consécutifs pour tous les instants du temps. Un concept développé plus tard est celui des convois proposé par Jeung *et al.* [34]. Un convoi correspond à un groupe contenant au moins m observations connectées par la densité par rapport à la distance ϵ et au cours de k instants consécutifs. D'autres critères ont été créés pour définir des types de trajectoires plus complexes comme les troupeaux, les essaims ou les rassemblements [34–37]. Cependant, des contraintes supplémentaires sont introduites

à chaque fois pour modéliser ces nouveaux ensembles de trajectoires, ce qui complexifie la paramétrisation des algorithmes et l'identification des groupes.

Une des dernières approches basée sur la densité locale de l'espace des caractéristiques, récemment proposée par Rodriguez et Laio [38], est le groupement par recherche rapide des pics de densité (DP). L'approche DP définit les centres des groupes comme étant des observations entourées par des voisins de densité locale moindre et éloignés des autres pics de densité. Afin de détecter les centres des groupes, l'algorithme utilise une distance de coupure d_c pour définir le voisinage local des observations. Les densités $\{\rho_i\}_{i=1...N}$ de chaque observation $\{\mathbf{x}_i\}_{i=1...N}$ sont ensuite déterminées en calculant :

$$\rho_i = \sum_j \chi(d_{ij} - d_c) \tag{3.6}$$

avec $\chi(u)=1$ si $u<0$ et $\chi(u)=0$ sinon. d_{ij} est la distance entre les observations $\mathbf{x}_i$ et $\mathbf{x}_j$.

Pour chaque $\mathbf{x}_i$ sa distance δ_i à l'observation de densité supérieure la plus proche est déterminée. Les centres des groupes sont initialisés par rapport à la valeur de $\gamma_i = \rho_i\delta_i$. En choisissant les K observations maximisant γ_i, cela permet de choisir les K maximums locaux respectivement à d_c. Une fois que les centres des groupes on été trouvés, chaque $\mathbf{x}_i$ est assigné au même groupe que son plus proche voisin de densité supérieure. Les points isolés, respectivement à d_c, sont quant à eux ignorés. Néanmoins, le choix du nombre de groupes reste à la charge de l'utilisateur.

La grande force des approches basées sur la densité comparées à celles présentées jusqu'à maintenant est qu'elles ne cherchent pas à expliquer toutes les données. Ce sont des méthodes capables de reconnaître des points isolés ou aberrants par rapport à un certain critère et de ne pas les prendre en compte dans le partitionnement des observations. De plus, aucune contrainte sur la forme des groupes dans l'espace des caractéristiques n'est imposée et la dimensionnalité des données n'est pas un problème en soi car ce sont les distances qui sont prises en compte, *e.g.* les auteurs ont retrouvé correctement avec DP 16 clusters dans un espace à 256 dimensions. Il est cependant important que la distance choisie retranscrive de manière adéquate les (dis)similarités existantes entre les observations. Par ailleurs, DP a été adaptée au groupement de séries temporelles par Begum *et al.* [13] en utilisant la DTW comme mesure de similarité.

Néanmoins, il reste deux points importants qui causent préjudice à cette méthode. Le premier est le choix de la distance de coupure d_c qui influe sur le résultat du groupement. Les auteurs disent que pour des jeux de données suffisamment grands les analyses sont robustes au choix de d_c, mais aucune valeur ne vient quantifier les termes "suffisamment grand". De plus, il est nécessaire de connaître a priori le nombre de groupes à retrouver, même si une technique est donnée pour guider le choix, cette procédure nécessite un choix humain. Ceci peut poser un problème sur des données réelles si aucune connaissance a priori n'est disponible ou si les données sont trop complexes pour évaluer le nombre de groupes. Enfin, nous nous interrogeons aussi sur l'efficacité d'algorithmes non itératifs en

présence de bruit pouvant partiellement mélanger des groupes initialement distincts.

3.6 Conclusion

Dans ce chapitre, nous avons abordé les grandes familles de groupement qui ont non seulement été utilisées pour le traitement de données statiques mais aussi pour le groupement de séries temporelles.

Bien que les algorithmes hiérarchiques puissent être déterministes ce n'est pas le cas des k-means qui s'initialisent classiquement de manière aléatoire, ce qui peut rendre leurs résultats non reproductibles. De plus, la reproductibilité de l'approche spectrale dépend de l'algorithme de groupement utilisé, ce qui ne permet pas de considérer les résultats obtenus comme déterministes. Dans chacun de ces cas, la connaissance du nombre de groupes à retrouver peut être problématique. En effet, les données réelles peuvent s'avérer complexes et l'information a priori n'est pas forcément accessible. Tout l'intérêt de la méthode de groupement que nous voulons utiliser est qu'elle identifie des groupes de séries temporelles similaires sans avoir de connaissance a priori sur leur nombre ou leur constitution.

C'est d'ailleurs l'avantage des méthodes basées sur la densité. Il n'y a pas besoin de spécifier un nombre de groupes à retrouver. Les seuls pré-requis à l'utilisation de DBSCAN sont de définir une contrainte de voisinage et un nombre minimal de voisins dans ce périmètre. Quant à l'adaptation de la méthode DP, il n'y a besoin de spécifier qu'une contrainte de voisinage. Cela fait peu de paramètres à régler, ce qui rend ces méthodes faciles d'utilisation. De plus, elles ne cherchent pas à "comprendre toutes les données", *i.e.* certains groupes peuvent ne comporter qu'un élément car il n'y a aucune obligation de tous les associer à un nombre de groupes prédéfinis : c'est la densité locale des éléments dans l'espace des caractéristiques qui prévaut. Cependant la contrainte de voisinage n'est pas intuitive à régler car dans le cas des séries temporelles elle représente un pourcentage de la distance, ce qui n'est pas du tout intuitif à cause de la haute dimensionnalité des séries temporelles. Le fait que ces méthodes ne soient pas itératives peut aussi causer un problème dans notre cas. En effet, ne pas affiner itérativement les groupements risque de compromettre l'identification précise des données réelles en présence de bruit. L'aspect discriminant d'algorithmes non itératifs alors que les différents groupes peuvent être partiellement mélangés est plus incertain.

Par conséquent, l'utilisation d'un algorithme de groupement itératif basé sur la densité nous semble plus approprié pour le groupement de séries temporelles réelles. C'est pourquoi nous voulons adapter mean-shift (M-S), une méthode basée sur l'estimation itérative des maxima locaux de la densité des observations, au groupement de séries temporelles.

Chapitre 4

Mean shift

4.1 Introduction

La méthode mean-shift (M-S) a été introduite par Fukunaga (1975) [39] mais n'a été popularisée que par Cheng en 1995 [40]. C'est ensuite Comaniciu [41] qui a utilisé mean-shift dans le contexte du filtrage et de la segmentation d'images.

A l'origine, cette méthode permet de faire converger des observations grâce à une procédure itérative vers les maxima locaux (les modes) de leur densité de probabilité sous-jacente. Les valeurs des modes sont attribuées aux observations après convergence. Néanmoins, cela décrit un processus de filtrage et non de groupement. Une fois que les modes de la densité de probabilité sont atteints, les valeurs des observations similaires sont très proches. Ainsi, l'étape de groupement devient triviale, elle sera détaillée en section 7.2.3. Nous présentons maintenant le formalisme mean-shift.

4.2 Formalisme

4.2.1 Estimation du gradient de la densité de probabilité

M-S est une méthode permettant de faire converger de manière itérative des observations $\mathbf{x}_i$ vers les modes de leur densité de probabilité f. Pour déplacer les observations vers les modes, il est nécessaire de calculer des estimation locales du gradient de la densité de probabilité $\hat{\nabla} f(\mathbf{x})$. Ces estimations locales, aux points $\mathbf{x}$, du gradient de la densité de probabilité f des observations $\mathbf{x}_i$, sont basées sur l'estimation non paramétrique de Parzen

(Parzen 1962 [43]) et s'expriment par :

$$\hat{\nabla} f(\mathbf{x}) \equiv \nabla \hat{f}(\mathbf{x}) = \frac{1}{n} \sum_{i=1}^{n} \nabla K_{\mathbf{H}}(\mathbf{x} - \mathbf{x}_i) \tag{4.1}$$

avec n le nombre d'observations, K une fonction appelée *noyau* qui est une fonction de densité de probabilité permettant d'en faire l'estimation locale, et $\mathbf{H}$ une matrice échelle carrée, symétrique et définie positive. En définissant $\mathbf{H}$ comme une matrice diagonale, les dimensions des observations sont normalisées indépendamment. Cependant rien n'oblige à choisir un telle matrice échelle, on peut imaginer utiliser des valeurs spécifiques $\mathrm{H}_{ij} \neq 0$ afin de combiner certaines caractéristiques entre elles. Les noyaux K sont des fonctions de densité de probabilité centrées et éventuellement limitées à un support borné définies de $\mathbb{R}^p$ dans $[0; 1]$, p étant la dimension des observations $\mathbf{x}$. Ce sont des fonctions positives, de moyenne nulle et dont les intégrales valent 1. K peut donc s'exprimer :

$$K_{\mathbf{H}}(\mathbf{x}) = |\mathbf{H}|^{-1/2}.K\Big(\mathbf{H}^{-1/2}\mathbf{x}\Big) \tag{4.2}$$

Les noyaux K sont généralement des fonctions sphériques. Par conséquent, ils peuvent être représentés par des noyaux monodimensionnels $k(x), x \geq 0$, appelés *profils*. Ainsi, un noyau peut être défini en fonction de son profil k par :

$$K(\mathbf{x}) = C.k(\mathbf{x}^T\mathbf{x}) \tag{4.3}$$

avec C la constante de normalisation du noyau. L'équation du gradient de l'estimation de la densité de probabilité (4.1) peut alors s'écrire :

$$\nabla \hat{f}(\mathbf{x}) = \frac{2.C.\mathbf{H}^{-1}}{n.|\mathbf{H}|^{1/2}} \sum_{i=1}^{n} (\mathbf{x} - \mathbf{x}_i).k'\Big((\mathbf{x} - \mathbf{x}_i)^T\mathbf{H}^{-1}(\mathbf{x} - \mathbf{x}_i)\Big) \tag{4.4}$$

avec $k'(x)$ la dérivée du profil k par rapport à x. En définissant $g(x) = -k'(x)$, et après quelques manipulations, cette expression devient :

$$\nabla \hat{f}(\mathbf{x}) = \frac{2.C.\mathbf{H}^{-1}}{n.|\mathbf{H}|^{1/2}} \sum_{i=1}^{n} g\Big(d^2(\mathbf{x}, \mathbf{x}_i, \mathbf{H})\Big) . \left[\frac{\sum_{i=1}^{n} g(d^2(\mathbf{x}, \mathbf{x}_i, \mathbf{H})) .\mathbf{x}_i}{\sum_{i=1}^{n} g(d^2(\mathbf{x}, \mathbf{x}_i, \mathbf{H}))} - \mathbf{x} \right] \tag{4.5}$$

avec $d(\mathbf{x}, \mathbf{x}_i, \mathbf{H}) = (\mathbf{x} - \mathbf{x}_i)^T\mathbf{H}^{-1}(\mathbf{x} - \mathbf{x}_i)$ la "distance euclidienne normalisée". Elle est souvent appelée distance de Mahalanobis dans le contexte mean-shift, bien que $\mathbf{H}$ ne soit pas une matrice de covariance ni $\mathbf{x}$ une moyenne. Le premier terme de cette équation est proportionnel à l'estimation de densité en $\mathbf{x}$. Le terme entre crochets est appelé vecteur mean-shift. Il décrit le déplacement dans la direction du gradient de la densité et exprimé sous une forme itérative, permettra de converger vers le mode local.

4.2.2 Méthode mean-shift

M-S considère chaque observation comme un vecteur de caractéristiques. En fonction des applications il peut s'agir de positions spatiales, de couleurs... Pour chaque observation, M-S définit un voisinage et calcule le barycentre des observations qu'il inclut. Les mesures contenues dans le vecteur de caractéristiques sont actualisées avec celles du barycentre de manière itérative, jusqu'à convergence aux modes de la densité de probabilité (Fig. 4.1). Celle-ci a été démontrée par Cheng [40] et Fashing [42].

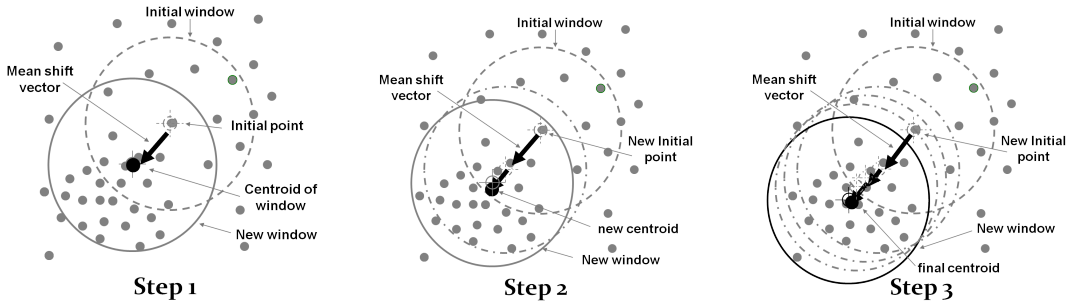


FIGURE 4.1 – Illustration, pour une observation donnée, du principe de la convergence M-S vers un mode.
source : http ://theses.insa-lyon.fr/publication/2012ISAL0030/these.pdf

Le problème M-S, celui de la remontée des observations vers les modes dans le sens du gradient de la densité, peut être formulé comme :

$$\nabla \hat{f}(\mathbf{x}) = 0 \tag{4.6}$$

Fukunaga [39] a proposé un algorithme itératif utilisant un fenêtrage de Parzen pour résoudre (4.6) en s'appuyant sur l'expression donnée en (4.5). Le vecteur mean-shift $m_i(\mathbf{x})$ des observations y est exprimé, permettant de définir l'équation de mise à jour des observations pour chaque itération k :

$$\begin{aligned} m_i(\mathbf{x}) = \mathbf{x}_i^{[k+1]} - \mathbf{x}_i^{[k]} &= \frac{\sum_{j=1}^{n} g\Big(d^2\Big(\mathbf{x}_i^{[k]}, \mathbf{x}_j^{[k]}, \mathbf{H}\Big)\Big) \cdot \mathbf{x}_j^{[k]}}{\sum_{j=1}^{n} g\Big(d^2\Big(\mathbf{x}_i^{[k]}, \mathbf{x}_j^{[k]}, \mathbf{H}\Big)\Big)} - \mathbf{x}_i^{[k]} \\ \mathbf{x}_i^{[k+1]} &= \frac{\sum_{j=1}^{n} g\Big(d^2\Big(\mathbf{x}_i^{[k]}, \mathbf{x}_j^{[k]}, \mathbf{H}\Big)\Big) \cdot \mathbf{x}_j^{[k]}}{\sum_{j=1}^{n} g\Big(d^2\Big(\mathbf{x}_i^{[k]}, \mathbf{x}_j^{[k]}, \mathbf{H}\Big)\Big)} \end{aligned} \tag{4.7}$$

avec un ensemble d'observations $\mathbf{X} = \{\mathbf{x}_i\}_{i=1\ldots n}$ et où $\mathbf{x}_i^{[k+1]}$ est la mise à jour de $\mathbf{x}_i^{[k]}$ à une position plus dense dans l'espace des caractéristiques.

La fonction $g : \mathbb{R} \to [0; 1]$ est une fonction permettant de pondérer la distance d et valant 1 pour $d(\cdot) = 0$. Si elle est définie décroissante elle permet d'accroître l'importance donnée aux points proches de l'observation $\mathbf{x}_i$ et à un certain point, de ne pas prendre en

compte les $\mathbf{x}_j$ trop éloignés de l'observation $\mathbf{x}_i$ à mettre à jour.

Il ne faut pas oublier que mean-shift est un algorithme de filtrage, par conséquent il faut grouper les observations ayant convergé pour obtenir les groupes. Comme énoncé précédemment, cette étape relativement simple est abordée en section 7.2.3.

L'approche mean-shift comprend tous les avantages déjà énoncés dans les sections précédentes des méthodes basées sur la densité et en plus, ne nécessite pas de connaître à l'avance ou de spécifier le nombre de groupes à trouver. Cette approche est aussi très adaptée au traitement d'images de part sa capacité à pouvoir traiter de manière indépendante les caractéristiques spatiales des pixels et les mesures leur étant associées. Dans ce cas, les noyaux utilisés pour pondérer les caractéristiques sur les différents supports peuvent être différents et l'équation 4.7 s'exprime désormais avec un produit de noyaux comme formalisé par Grenier [43] :

$$\mathbf{x}_i^{[\mathrm{k}+1]} = \frac{\sum_{j=1}^{n} \prod_{c=1}^{\mathrm{C}} g_c\Big(d_c^2\Big(\mathbf{x}_{c,i}^{[\mathrm{k}]}, \mathbf{x}_{c,j}^{[\mathrm{k}]}, \mathbf{H}_c\Big)\Big) \cdot \mathbf{x}_j^{[\mathrm{k}]}}{\sum_{j=1}^{n} \prod_{c=1}^{\mathrm{C}} g_c\Big(d_c^2\Big(\mathbf{x}_{c,i}^{[\mathrm{k}]}, \mathbf{x}_{c,j}^{[\mathrm{k}]}, \mathbf{H}_c\Big)\Big)} \tag{4.8}$$

avec g_c, d_c, $\mathbf{H}_c$ et $\mathbf{x}_c$ respectivement les noyaux, distances, matrices échelles et les mesures des échantillons associés à chaque support C. Classiquement, $C = 2$ en traitement d'images. Un noyau est utilisé pour le voisinage spatial et un noyau est utilisé pour le voisinage des mesures dans l'espace amplitude (intensité, couleurs...). La formulation classique de M-S en traitement d'images est donc la suivante :

$$\mathbf{x}_i^{[\mathrm{k}+1]} = \frac{\sum_{j=1}^{n} g_s\Big(d_s^2\Big(\mathbf{x}_{s,i}^{[\mathrm{k}]}, \mathbf{x}_{s,j}^{[\mathrm{k}]}, \mathbf{H}_s\Big)\Big) . g_r\Big(d_r^2\Big(\mathbf{x}_{r,i}^{[\mathrm{k}]}, \mathbf{x}_{r,j}^{[\mathrm{k}]}, \mathbf{H}_r\Big)\Big) . \mathbf{x}_j^{[\mathrm{k}]}}{\sum_{j=1}^{n} g_s\Big(d_s^2\Big(\mathbf{x}_{s,i}^{[\mathrm{k}]}, \mathbf{x}_{s,j}^{[\mathrm{k}]}, \mathbf{H}_s\Big)\Big) . g_r\Big(d_r^2\Big(\mathbf{x}_{r,i}^{[\mathrm{k}]}, \mathbf{x}_{r,j}^{[\mathrm{k}]}, \mathbf{H}_r\Big)\Big)} \tag{4.9}$$

avec $g_{s/r}$, $d_{s/r}$, $\mathbf{H}_{s/r}$ et $\mathbf{x}_{s/r}$ la fonction de pondération, la distance, la matrice échelle et les mesures associées respectivement aux supports spatial et amplitude.

La méthode M-S est sensible aux choix des paramètres d'échelle $\mathbf{H}_{s/r}$, il faut donc être vigilant dans leur choix pour ne pas compromettre les résultats. Cela reste assez simple avec des mesures monodimensionnelles, mais peut s'avérer plus compliqué avec des mesures de dimensionnalité élevée. De plus, l'efficacité de mean-shift risque de diminuer avec une trop grande augmentation du nombre de dimensions. Si les mesures deviennent éparses dans l'espace des caractéristiques elles pourraient toutes être des maxima locaux de densité, dans le pire des cas. Certaines études ont été menées sur la sensibilité du réglage des paramètres d'échelles ainsi que sur la détermination automatique de leurs valeurs optimales pour l'utilisation de M-S [44–48]. Néanmoins, cela représente un sujet d'étude à part entière, les solutions proposées donnent de bons résultats sur des données de faible dimensionnalité, ce qui n'est pas le cas avec les données temporelles telles que nous les définissons dans la suite du manuscrit. C'est pourquoi nous étudions empiriquement, en section 7.3.3, les

résultats obtenus avec le formalisme que nous proposons pour le traitement de données temporelles avec de multiples combinaisons des paramètres d'échelles.

A notre connaissance, Feng *et al.* [49] ont été les premiers à proposer une extension du formalisme mean-shift aux domaines espace/amplitude/temps dans le cadre du filtrage de séquences vidéo. Les techniques de filtrage spatio-temporel peuvent être séparées en deux catégories. Les approches causales [50, 51] n'utilisent que l'information passée dans le processus de filtrage. Au contraire, les approches omniscientes [46, 52–54] utilisent en totalité ou en partie l'information passée et future pour traiter les données. Mean-shift a aussi été appliqué pour le traitement de séries temporelles [55–57]. Néanmoins, il à été utilisé à chaque fois dans sa forme classique. Les séries temporelles sont considérées comme des vecteurs à T dimensions utilisés comme mesures multidimensionnelles dans l'espace amplitude, sans considérer le temps. Dans ces travaux, M-S est aussi utilisé sur des ensembles de statistiques calculées sur les séries temporelles.

4.3 Conclusion

Dans ce chapitre nous avons présenté les fondements ainsi que le formalisme mean-shift. Cette méthode comporte les avantages des méthodes basées sur la densité et nous avons rappelé que son formalisme est aussi très bien adapté au traitement d'images. Bien que M-S ait déjà été utilisé dans des études traitant des données temporelles, aucun formalisme permettant d'adapter cette méthode au groupement de séries temporelles n'a encore été proposé. C'est ce que nous proposons dans la partie III.

Chapitre 5

Conclusion

Dans cette partie, nous avons tout d'abord abordé au chapitre 2 les différentes mesures de comparaison des séries temporelles. Nous avons balayé les grandes familles de mesures, des normes de type L_p à des mesures plus complexes invariantes aux distorsions temporelles comme la mesure DTW.

Dans un second temps, nous avons présenté les principaux algorithmes de groupement, classiquement utilisés pour le groupement de données statiques mais étendus aux séries temporelles. Ces algorithmes ont par conséquent tous été adaptés pour utiliser une mesure appartenant à une des familles présentées au chapitre 3, excepté M-S à notre connaissance.

Cette première partie nous a permis de faire un état de l'art sur le groupement de séries temporelles et de montrer le potentiel des approches basées sur la densité pour traiter des signaux temporels réels issus d'images. Nous proposons dans la partie suivante d'étendre le formalisme mean-shift afin de permettre le filtrage et le groupement de séries temporelles.

III Contributions méthodologiques

Chapitre 6

Introduction

Dans cette partie, nous présentons trois contributions méthodologiques permettant de grouper des séries temporelles extraites de séquences d'images. Elles étendent le formalisme mean-shift qui différencie les caractéristiques appartenant au domaine spatial et celles appartenant au domaine des amplitudes. De plus, cette méthode ne nécessite aucun a priori sur la composition ou le nombre de groupes à identifier.

Notre première contribution est une méthode combinant filtrage et groupement de données spatio-temporelles multidimensionnelles. Cette approche adapte le formalisme mean-shift pour faire converger les évolutions temporelles de mesures associées à des échantillons dans l'espace amplitude. Pour ce faire, nous avons d'une part modifié le domaine joint spatial/amplitude afin de prendre en compte les évolutions temporelles des mesures et d'autre part, intégré au formalisme une contrainte sur la similarité des trajectoires des échantillons dans l'espace amplitude avec l'utilisation de la norme infini.

Notre deuxième contribution porte sur une méthode de groupement hiérarchique permettant de grouper des séries temporelles similaires et déphasées dans le temps. Une contrainte sur l'écart maximal entre les appariements optimaux retournés par la DTW, toujours définie par la norme infini, est utilisée pour déterminer le seuil d'acceptation d'un échantillon à un groupe.

Notre troisième contributions propose d'étendre le formalisme mean-shift en exprimant la valeur temporelle des échantillons de manière explicite et en la prenant en compte dans le processus de filtrage. Ce nouveau formalisme permet d'unifier, grâce à l'ajout d'un noyau temporel, les filtrages M-S de type causal et omniscient déjà existants.

Chapitre 7

Mean-shift spatio-temporel : STM-S

7.1 Introduction

Dans cette partie, nous adaptons le formalisme M-S au filtrage de séries temporelles (section 7.2.1). Celui-ci sera étendu aux mesures d'amplitudes multidimensionnelles (section 7.2.2) et intégrera une étape de groupement (section 7.2.3) successive ou conjointe au processus itératif de filtrage (section 7.2.4). Pour terminer, nous présenterons en section 7.3 les données simulées ainsi que notre démarche et les résultats obtenus pour la validation de notre approche.

7.2 Méthode

7.2.1 Formalisme

L'approche STM-S est basée sur la procédure M-S présentée en section 4.2. Dans un premier temps, les n échantillons $\{\mathbf{x}_i\}_{i=1\ldots n}$ à traiter sont définis par :

$$\mathbf{x}_i = \left[\mathbf{x}_{\mathrm{s},i}^{'}\ \mathbf{x}_{\mathrm{t},i}^{'}\right]^{'} \in \mathbf{X} \qquad \text{avec} \qquad \begin{array}{l} \mathbf{x}_{\mathrm{s},i} \in \mathbb{R}^{\mathrm{S}} : \text{ support spatial} \\ \mathbf{x}_{\mathrm{t},i} \in \mathbb{R}^{\mathrm{T}} : \text{ support amplitude} \\ i = 1, \ldots, n : \text{ identifiant de l'échantillon} \end{array} \tag{7.1}$$

Un échantillon $\mathbf{x}_i$ est donc constitué d'une position $\mathbf{x}_{\mathrm{s},i}$, à valeurs dans $\mathbb{R}^{\mathrm{S}}$, et d'une série de mesures $\mathbf{x}_{\mathrm{t},i}$ observées à T instants, à valeurs dans $\mathbb{R}^{\mathrm{T}}$. Ceci étant l'équation régissant le filtrage STM-S d'un échantillon, mettant à jour sa valeur entre les itérations k et $\mathrm{k}+1$,

est notée :

$$\mathbf{x}_i^{[\mathrm{k}+1]} = \frac{\sum_{j=1}^{n} Sp_{i,j}\left(\mathbf{x}_{\mathrm{s},i}^{[\mathrm{k}]}, \mathbf{x}_{\mathrm{s},j}^{[\mathrm{k}]}\right) \cdot Ra_{i,j}\left(\mathbf{x}_{\mathrm{t},i}^{[\mathrm{k}]}, \mathbf{x}_{\mathrm{t},j}^{[\mathrm{k}]}\right) \cdot \mathbf{x}_j^{[\mathrm{k}]}}{\sum_{j=1}^{n} Sp_{i,j}\left(\mathbf{x}_{\mathrm{s},i}^{[\mathrm{k}]}, \mathbf{x}_{\mathrm{s},j}^{[\mathrm{k}]}\right) \cdot Ra_{i,j}\left(\mathbf{x}_{\mathrm{t},i}^{[\mathrm{k}]}, \mathbf{x}_{\mathrm{t},j}^{[\mathrm{k}]}\right)} \tag{7.2}$$

où $Sp_{i,j}(\cdot)$ et $Ra_{i,j}(\cdot)$ correspondent respectivement aux distances pondérées, entre deux échantillons $\mathbf{x}_i$ et $\mathbf{x}_j$, utilisées avec les supports spatial et amplitude :

$$Sp_{i,j}\left(\mathbf{x}_{\mathrm{s},i}^{[\mathrm{k}]}, \mathbf{x}_{\mathrm{s},j}^{[\mathrm{k}]}\right) = g_{\mathrm{s}}\Big(d_{\mathrm{s}}^2\left(\mathbf{x}_{\mathrm{s},i}^{[\mathrm{k}]}, \mathbf{x}_{\mathrm{s},j}^{[\mathrm{k}]}, \mathbf{H}_{\mathrm{s}}\right)\Big) \tag{7.3}$$

$$Ra_{i,j}\left(\mathbf{x}_{\mathrm{t},i}^{[\mathrm{k}]}, \mathbf{x}_{\mathrm{t},j}^{[\mathrm{k}]}\right) = g_{\mathrm{r}}\Big(d_{\mathrm{r}}^2\left(\mathbf{x}_{\mathrm{t},i}^{[\mathrm{k}]}, \mathbf{x}_{\mathrm{t},j}^{[\mathrm{k}]}, \mathbf{H}_{\mathrm{r}}\right)\Big) \tag{7.4}$$

D'une part, la distance séparant deux positions $\mathbf{u}_{\mathrm{s}}$ et $\mathbf{v}_{\mathrm{s}}$ est calculée via la distance de Mahalanobis $d_{\mathrm{s}}\left(\mathbf{u}_{\mathrm{s}}, \mathbf{v}_{\mathrm{s}}, \mathbf{H}_{\mathrm{s}}\right)$. $\mathbf{H}_{\mathrm{s}}$ étant une matrice échelle de taille S×S, diagonale et définie positive telle que :

$$\mathbf{H}_{\mathrm{s}} = h_{\mathrm{s}} \cdot \mathbf{I} \tag{7.5}$$

avec h_{s} un scalaire représentant la valeur de l'échelle spatiale et $\mathbf{I}$ la matrice identité. D'autre part, la distance utilisée pour calculer l'éloignement entre deux mesures $\mathbf{u}_{\mathrm{r}}$ et $\mathbf{v}_{\mathrm{r}}$ est la norme infini :

$$d_{\mathrm{r}}\left(\mathbf{u}_{\mathrm{t}}, \mathbf{v}_{\mathrm{t}}, \mathbf{H}_{\mathrm{r}}\right) = \| \mathbf{H}_{\mathrm{r}}^{-\frac{1}{2}} \left(\mathbf{u}_{\mathrm{t}} - \mathbf{v}_{\mathrm{t}}\right) \|_{\infty} \tag{7.6}$$

Celle-ci transcrit l'éloignement maximal, temps à temps, entre deux évolutions de mesures. $\mathbf{H}_{\mathrm{r}}$ est une matrice échelle de taille T×T appliquée aux mesures, diagonale et définie positive telle que :

$$\mathbf{H}_{\mathrm{r}} = h_{\mathrm{r}} \cdot \mathbf{I} \tag{7.7}$$

avec h_{r} un scalaire représentant la valeur de l'échelle dans l'espace amplitude et $\mathbf{I}$ la matrice identité. Le cas où $\mathbf{H}_{\mathrm{r}}$ est définie comme une matrice diagonale est un cas particulier où les mesures sont normalisées de manière indépendante (cf. section 4.2). Les distances calculées sur chaque support sont ensuite pondérées par une fonction $g(\cdot)$ déterminant la participation ou non d'un échantillon candidat $\mathbf{x}_j$ au filtrage d'un échantillon référence $\mathbf{x}_i$ en fonction de leur éloignement :

$$g_{\mathrm{s}}\Big(d_{\mathrm{s}}^2\left(\cdot\right)\Big) = g_{\mathrm{r}}\Big(d_{\mathrm{r}}^2\left(\cdot\right)\Big) = \begin{cases} 1 & \text{si } d_{\mathrm{s}}^2\left(\cdot\right),\ d_{\mathrm{r}}^2\left(\cdot\right) \leq 1 \\ 0 & \text{sinon} \end{cases} \tag{7.8}$$

Pour rappel, le choix de définir $g(\cdot)$ comme une fonction porte est un cas particulier d'utilisation. D'autres fonctions de pondération pourraient être utilisées (cf. section 4.2). Si l'éloignement sur un des supports dépasse la valeur fixée par la matrice échelle associée, alors l'échantillon candidat sera exclu du filtrage de l'échantillon référence pour l'itération courante. Par conséquent, la combinaison des définitions (7.3), (7.4) et (7.8) dans la formule générale du filtrage STM-S d'un échantillon (7.2) mène à uniquement considérer les échantillons qui lui seront proches spatialement au regard de l'échelle $\mathbf{H}_{\mathrm{s}}$, et dont les

évolutions de mesures lui resteront suffisamment proches, temps à temps, au regard de l'échelle $\mathbf{H}_{\mathrm{r}}$.

La figure 7.1 illustre le principe de sélection des échantillons décrit par les équations précédentes. L'échantillon rouge est inclu dans le voisinage spatial de l'échantillon référence (en bleu), cependant l'évolution temporelle de sa mesure en amplitude n'est pas strictement inclue dans la "gaine temporelle" définie par les pointillés bleus (Fig. 7.1) autour de l'évolution référence (7.6). Par conséquent, le candidat rouge ne sera pas retenu pour la mise à jour de l'échantillon référence (7.2). L'échantillon vert est quant à lui suffisamment proche spatialement et temporellement de l'échantillon référence pour participer à la mise à jour de l'échantillon de référence.

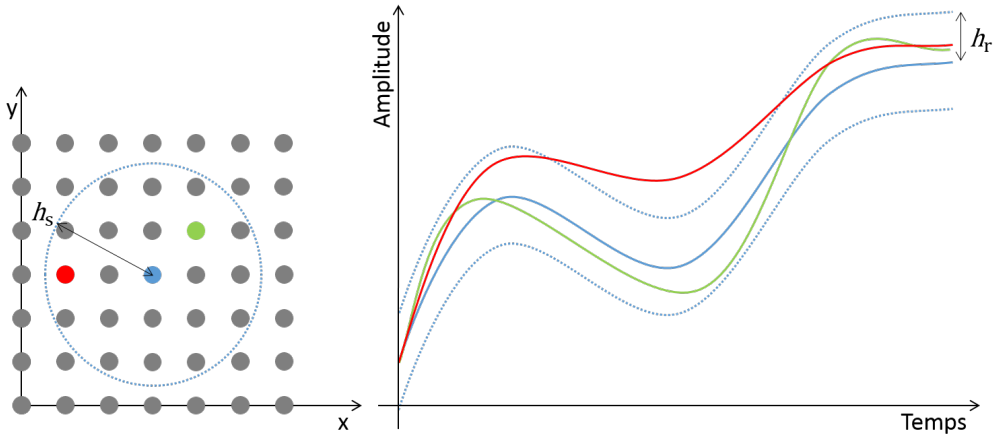


FIGURE 7.1 – A gauche : mesures spatiales des échantillons extraits d'une image 2D à la première itération de filtrage (k = 0). A droite : évolutions temporelles des mesures dans le domaine amplitude de trois échantillons, les évolutions sont associées aux mesures spatiales de même couleur. Le bleu identifie l'échantillon référence, tandis que le vert et le rouge identifient deux des échantillons candidats. Les lignes bleues en pointillés représentent les limites des voisinages spatial et temporel. Les mesures associées aux échantillons candidats doivent être incluses strictement entre chacune de ces limites pour être prises en compte dans (7.2).

L'algorithme 2 décrit la mise en œuvre des équations précédentes pour réaliser le filtrage spatio-temporel des échantillons. Cet algorithme itératif est inspiré de l'approche mean-shift présentée section 4.2.

7.2.2 Extension du formalisme aux mesures multidimensionnelles dans le domaine des amplitudes

Le formalisme STM-S est aisément extensible au filtrage d'échantillons dont les mesures dans le domaine amplitude sont multicomposantes, *e.g.* séquence d'images couleurs constituées d'un plan rouge, d'un plan vert et d'un plan bleu. Les mesures $\mathbf{x}_{\mathrm{t},i}$ s'expriment désormais :

$$\mathbf{x}_{\mathrm{t},i} = \left[\left(\mathbf{x}_{\mathrm{t},i}^{\mathrm{C}_1} \right)' \ \ldots \ \left(\mathbf{x}_{\mathrm{t},i}^{\mathrm{C}_f} \right)' \right]' \tag{7.9}$$

Algorithm 2 Algorithme du filtrage STM-S

Require: h_s et h_r fixés par l'utilisateur
Require: $\mathbf{X} = \{\mathbf{x}_i\}_{i=1\ldots n}$ ensemble d'entrée des échantillons
1: $k \leftarrow 0$
2: $\mathbf{X}^{[0]} \leftarrow \mathbf{X}$
3: **repeat**
4: **for all** $\mathbf{x}_i \in \mathbf{X}^{[\mathrm{k}]}$ **do**
5: Calcul de $\mathbf{x}_i^{[\mathrm{k+1}]}$ avec (7.2)
6: **end for**
7: $k \leftarrow k + 1$
8: **until** $|\mathbf{X}^{[\mathrm{k}]} - \mathbf{X}^{[\mathrm{k\text{-}1}]}| < \epsilon$
9: $\hat{\mathbf{X}} \leftarrow \mathbf{X}^{[\mathrm{k}]}$
10: **return** $\hat{\mathbf{X}}$

Chacune des mesures C_f, avec $f = 1, \ldots, \mathrm{F}$, est réalisée à T instants. Le vecteur de mesures $\mathbf{x}_{\mathrm{t},i}$ est donc à valeurs dans $\mathbb{R}^{\mathrm{T}\times\mathrm{F}}$.
La distance $d_\mathrm{r}(\cdot)$ est conservée dans le cas multidimensionnel. La matrice d'échelle en amplitude $\mathbf{H}_\mathrm{r}$ de taille T×F × T×F est diagonale par blocs de taille T×T, ce qui permet de combiner ou non les mesures C_f. Par exemple, en considérant des mesures tridimensionnelles dans l'espace amplitude (C_1, C_2, C_3) et en combinant C_1 et C_2, la matrice échelle en amplitude devrait être définie comme :

$$\mathbf{H}_\mathrm{r} = \begin{bmatrix} \mathbf{H}_\mathrm{r}^{\mathrm{C}_1} & \mathbf{H}_\mathrm{r}^{\mathrm{C}_2} & \mathbf{Ø} \\ \mathbf{Ø} & \mathbf{Ø} & \mathbf{Ø} \\ \mathbf{Ø} & \mathbf{Ø} & \mathbf{H}_\mathrm{r}^{\mathrm{C}_3} \end{bmatrix} \quad \text{ou} \quad \mathbf{H}_\mathrm{r} = \begin{bmatrix} \mathbf{H}_\mathrm{r}^{\mathrm{C}_1} & \mathbf{Ø} & \mathbf{Ø} \\ \mathbf{H}_\mathrm{r}^{\mathrm{C}_2} & \mathbf{Ø} & \mathbf{Ø} \\ \mathbf{Ø} & \mathbf{Ø} & \mathbf{H}_\mathrm{r}^{\mathrm{C}_3} \end{bmatrix} \tag{7.10}$$

On peut noter que la forme diagonale par blocs permettra de considérer les mesures de manière indépendante.

7.2.3 Groupement des échantillons

La procédure M-S est par nature un processus de filtrage. Le nombre d'échantillons est conservé entre la première et la dernière itération, durant lesquelles l'algorithme fait converger les échantillons similaires vers le plus représentatif, à savoir les mesures du représentant moyen. De fait, il est nécessaire de grouper les échantillons similaires afin de pouvoir étudier les différentes classes d'évolution présentes dans les images ou pour faire de la segmentation.

Par conséquent, une étape de groupement est associée de manière conjointe ou postérieure au filtrage des échantillons. Ceci permet de mettre en évidence de manière non supervisée les différentes classes d'évolution, aux échelles spécifiées, et d'accélérer la convergence quand des groupements intermédiaires sont effectués à la fin de chaque itération. De plus, lorsque que les groupements sont effectués simultanément au filtrage le nombre d'échantillons à traiter s'en trouve diminué, ce qui augmente la vitesse de convergence de

l'algorithme. Le groupement simultané des échantillons aux itérations de filtrage ainsi que le gain apporté en temps de calcul seront discutés en section 7.2.4.

L'algorithme de groupement que nous proposons (cf. Algorithme 3) permet d'identifier les différents groupes après convergence des échantillons. Dans un premier temps les échantillons sont associés à leur mode de convergence, *i.e.* groupés avec tous les échantillons suffisamment proches d'eux au regard des échelles spatiale et temporelle. Deuxièmement, les groupes dont les distances inter-classes en amplitude sont inférieures à l'échelle fixée sont désignés comme candidats pour le groupement final. Ainsi les groupes avec des représentants dont les évolutions de mesures dans le domaine amplitude sont proches devraient ne former qu'une seule classe. Néanmoins ces groupes n'appartiennent pas forcément à la même région de l'image, c'est pourquoi les classes candidates sont regroupées si et seulement si leurs échantillons appartiennent à des régions connexes dans l'espace image initial. A chaque étape du groupement, le représentant de chaque classe est défini comme le barycentre des échantillons associés.

Nous procédons ainsi car après une convergence de type M-S, il est commun de voir apparaître des "artéfacts de régions" dans les images. Si les échelles spatiales sont sous-dimensionnées par rapport à la taille de régions homogènes à identifier, les échantillons leur appartenant vont converger vers des maxima de densité différents. Par conséquent les régions homogènes apparaîtront morcelées, *i.e.* le nombre de groupes sera surestimé après groupement des échantillons. L'algorithme de groupement proposé est présenté ci-dessous :

Algorithm 3 Groupement des échantillons après convergence de la procédure STM-S

Require: $\hat{\mathbf{X}} = \{\hat{\mathbf{x}}_i\}_{i=1\ldots n}$: Ensemble retourné par l'algorithme 2 (filtrage STM-S)
1: $\mathbf{G} \leftarrow$ Identifier chaque échantillon par son groupe de convergence
2: $\mathbf{G}^* \leftarrow$ Identifier de manière unique chaque groupe de $\mathbf{G}$
3: **for all** $\mathrm{g}_i^* \in \mathbf{G}^*$ **do**
4: $\quad\mathcal{N} \leftarrow$ Trouver les similaires à g_i^* dans le domaine amplitude relativement à $\mathbf{H}_\mathrm{r}$
5: $\quad\mathcal{S} \leftarrow$ Extraire les groupes connexes à g_i^* une fois les échantillons associés au(x) groupe(s) de l'ensemble $\mathcal{N}$ et à g_i^* repropagés sur la grille spatiale de la séquence avant filtrage
6: $\quad\mathbf{G} \leftarrow$ Regrouper les échantillons associés au(x) groupe(s) de l'ensemble $\mathcal{S}$ avec ceux de g_i^*
7: $\quad\mathbf{G}^* \leftarrow$ Supprimer le ou les identifiant(s) associé(s) au(x) groupe(s) de l'ensemble $\mathcal{S}$
8: **end for**
9: **return** $\mathbf{G}$

7.2.4 Groupement des échantillons conjoint au filtrage

Afin d'accélérer la convergence des échantillons, nous proposons un algorithme de groupement des échantillons, après chaque itération du filtrage STM-S. Ce groupement consiste à fusionner les échantillons très proches sur tous les supports à la fin de chaque itération, en faisant l'hypothèse qu'ils vont converger aux mêmes modes. Ainsi, seul le représentant

des échantillons participera au filtrage et le poids lui étant associé sera proportionnel au nombre d'échantillons représentés. L'équation (7.2) s'exprime désormais :

$$\mathbf{x}_i^{[\mathrm{k}+1]} = \frac{\sum_{j=1}^{n} Sp_{i,j}(\cdot)\,.Ra_{i,j}(\cdot)\,.w_j^{[\mathrm{k}]}.\mathbf{x}_j^{[\mathrm{k}]}}{\sum_{j=1}^{n} Sp_{i,j}(\cdot)\,.Ra_{i,j}(\cdot)\,.w_j^{[\mathrm{k}]}} \tag{7.11}$$

où $w_j^{[\mathrm{k}]}$ correspond au nombre d'échantillons représentés par l'échantillon j et où n correspond au nombre d'échantillons à l'itération k.

Les poids $w_i^{[\mathrm{k}+1]}$ sont actualisés après le calcul des $\mathbf{x}_i^{[\mathrm{k}+1]}$ en regroupant les échantillons très proches. Cette proximité est déduite des valeurs de $d_{\mathrm{s}}(\cdot)$ et $d_{\mathrm{r}}(\cdot)$ en utilisant des matrices échelles $\mathbf{H}_{\mathrm{s}}$ et $\mathbf{H}_{\mathrm{r}}$ très sélectives : 1000 fois inférieures aux échelles utilisées pour le filtrage. Il est possible de faire varier ce seuil, néanmoins nous avons validé ce choix empiriquement sur différents jeux de données. Dans tous les cas, cette paramétrisation permet d'obtenir un compromis satisfaisant entre rapidité de convergence et qualité des résultats. Afin de minimiser l'impact du groupement sur le processus mean-shift spatio-temporel, c'est le barycentre des échantillons à regrouper, pondéré par $w^{[\mathrm{k}]}$, qui les représentera à l'itération suivante. Le principe de convergence de STM-S avec groupement des échantillons conjoint au filtrage est illustré Fig. 7.2.

Considérant $\bar{k}$ le nombre moyen d'itérations pour atteindre la convergence, le coût algorithmique du filtrage mean-shift spatio-temporel est de l'ordre de $\Theta(\bar{k}n^2)$. Nous avons observé que l'algorithme de groupement permet de diminuer, en moyenne, le nombre d'échantillons de moitié à chaque itération ($n^{[\mathrm{k}+1]} \lesssim n^{[\mathrm{k}]}/2$). En partant de cette observation, calculons le nombre moyen d'échantillons traités au cours de cette nouvelle procédure STM-S :

$$\begin{aligned} n_{moy} &= n + \frac{n}{2} + \frac{n}{4} + \frac{n}{8} + \frac{n}{16} + \dots \\ &= n\sum_{\mathrm{k}=0}^{\bar{\mathrm{k}}} \frac{1}{2^{\mathrm{k}}} = n\sum_{\mathrm{k}=0}^{\bar{\mathrm{k}}} \left(\frac{1}{2}\right)^{\mathrm{k}} < 2n \end{aligned} \tag{7.12}$$

Par conséquent, le coût algorithmique moyen du filtrage combiné au groupement des échantillons est de l'ordre de $\Theta(4n^2)$. Le nombre moyen d'itérations n'ayant plus d'impact sur le coût algorithmique, le gain de temps pour faire converger les échantillons est potentiellement considérable.

Afin de valider le calcul théorique du gain de temps inhérent à l'algorithme de groupement, nous avons testé cette nouvelle procédure sur un jeu de données réelles composé de 669962 échantillons. Il s'agit d'un cerveau de patient atteint de sclérose en plaques acquis en trois dimensions en imagerie par résonance magnétique (Fig. 7.3). La diminution du nombre d'échantillons et l'évolution du temps passé à chaque itération de filtrage sont présentés Fig. 7.4. Nous remarquons Fig. 7.4(a) que malgré une baisse plus importante du nombre d'échantillons au cours des premières itérations, le facteur moyen de diminution

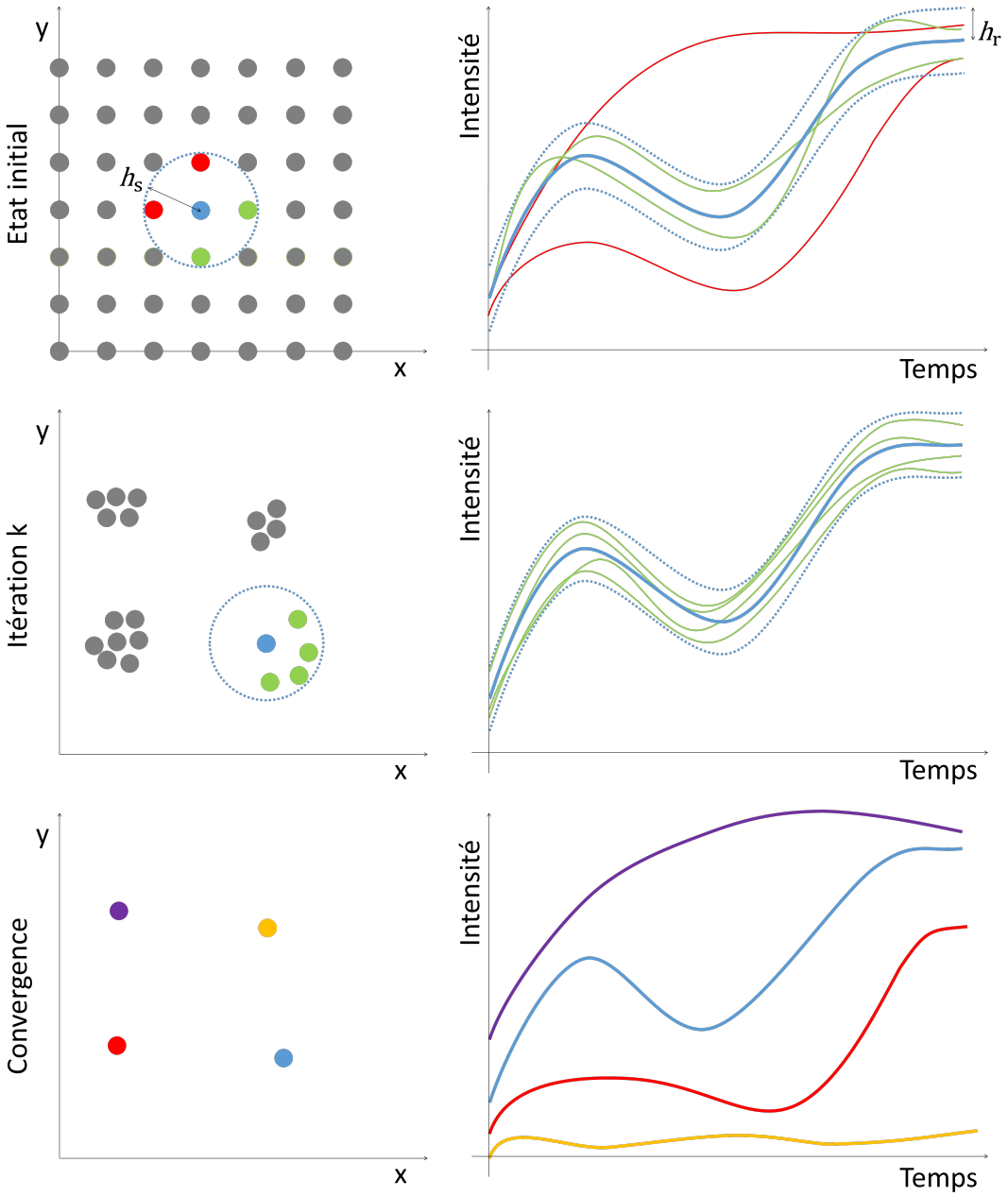
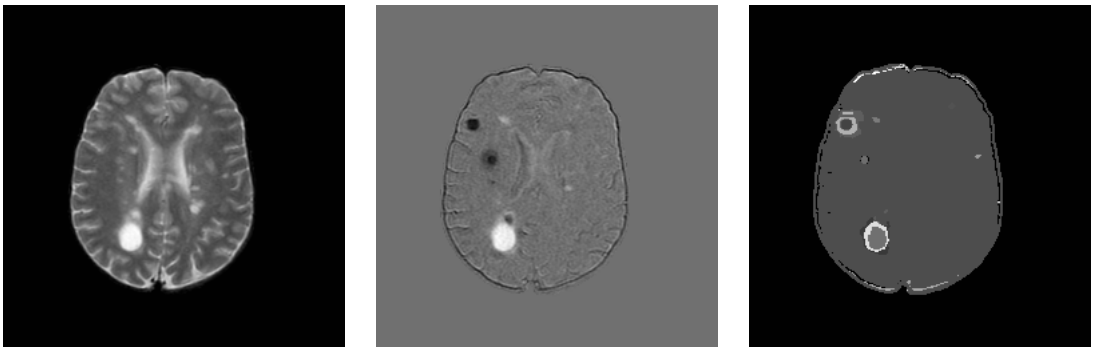


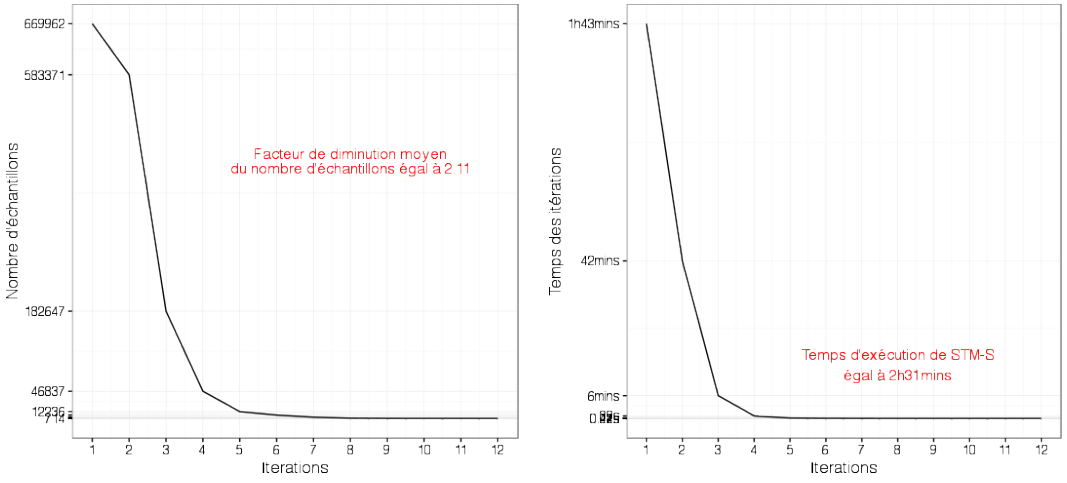
FIGURE 7.2 – Illustration de la convergence de STM-S. A gauche : support spatial. A droite : Support des amplitudes, évolution d'intensités. Les pointillés indiquent les voisinages sur les supports spatial et amplitude. L'observation bleue est celle utilisée comme référence à chaque itération. Les observations rouges sont les candidats pour mettre à jour la référence, mais qui ne sont pas à la fois compris dans le voisinage spatial et temporel de la référence (strictement entre les pointillées). Les observations vertes sont les candidats qui satisfont ces deux conditions et servent à calculer le barycentre pour mettre à jour l'échantillon référence. Le nombre d'observations diminue au fil des itérations en utilisant l'algorithme de groupement conjointement au filtrage.



(a) (b) (c)

FIGURE 7.3 – (a) Coupe d'une image IRM aquise en trois dimensions d'un patient atteint de sclérose en plaques. (b) Image de différence entre l'image en (a) et la première acquisition du suivi du patient. (c) Résultat de STM-S après groupement conjoint au filtrage en trois dimensions du volume montré en (b).

des échantillons entre deux itérations est égal à $2,11$, ce qui correspond à l'hypothèse faite pour le calcul de n_{moy} (7.12). Nous pouvons aussi voir Fig. 7.4(b) que le temps de calcul diminue très fortement sur les premières itérations. C'est un résultat logique car la complexité en temps de STM-S est quadratique. Par conséquent, les plus grosses diminutions de temps d'itérations sont celles ou le nombre d'échantillons a le plus diminué. Le temps total de STM-S avec l'algorithme de groupement est de 2h31mins tandis qu'il aurait été classiquement de 20h36mins (1h43mins $\times$ 12 itérations). Le facteur d'accélération de la convergence de STM-S par l'algorithme de groupement sur ces données est donc environ égal à 8. L'algorithme de groupement des échantillons associé au filtrage STM-S est détaillé dans l'algorithme 4.



(a) (b)

FIGURE 7.4 – (a) Nombre d'échantillons traités à chaque itération de STM-S avec l'algorithme de groupement. (b) Temps de calcul des itérations de STM-S avec l'algorithme de groupement.

Algorithm 4 Algorithme groupement des échantillons conjoint au filtrage STM-S

Require: $\mathbf{X} = \{\mathbf{x}_i\}_{i=1...n^{[0]}}$
Require: $\mathbf{W} = \{\mathbf{w}_i\}_{i=1...n^{[0]}}$
k $\leftarrow 0$, $d \leftarrow +\infty$
$\mathbf{X}^{[0]} \leftarrow \mathbf{X}$, $\mathbf{W}^{[0]} \leftarrow \mathbf{W}$
repeat
  for all $\mathbf{x}_i \in \mathbf{X}^{[\mathrm{k}]}$ **do**
    Calcul de $\mathbf{x}_i^{[\mathrm{k}+1]}$ avec (7.11)
  end for
  $d \leftarrow |\mathbf{X}^{[\mathrm{k}+1]} - \mathbf{X}^{[\mathrm{k}]}|$
  $\mathbf{M} \leftarrow$ tous les $\mathbf{x}_i^{[\mathrm{k}]} \in \mathbf{X}^{[\mathrm{k}]}$ qui peuvent être groupés
  repeat
    Extraire un $\mathbf{x}_i$ de $\mathbf{M}$
    $\mathbf{x}_i^{[\mathrm{k}+1]} \leftarrow$ Grouper $\mathbf{x}_i^{[\mathrm{k}+1]}$ avec ses voisins dans $\mathbf{M}$
    $\mathbf{W}^{[\mathrm{k}+1]} \leftarrow$ Mettre à jour la pondération de $\mathbf{x}_i^{[\mathrm{k}+1]}$
    Supprimer les voisins de $\mathbf{x}_i^{[\mathrm{k}+1]}$ des ensembles $\mathbf{X}^{[\mathrm{k}+1]}, \mathbf{W}^{[\mathrm{k}+1]}$ et $\mathbf{M}$
  until $\mathbf{M} = \emptyset$
  $k \leftarrow k + 1$
until $d < \epsilon$
$\hat{\mathbf{X}} \leftarrow \mathbf{X}^{[\mathrm{k}+1]}$
return $\hat{\mathbf{X}}$

7.3 Validation

Nous venons d'exposer le formalisme STM-S ainsi que l'algorithme de groupement appliqué pendant et/ou après convergence des échantillons. Dans un premier temps nous présentons les mesures et les données simulées permettant de valider la qualité, la robustesse et l'intérêt de la méthode proposée. Deuxièmement, nous évaluons le gain apporté par la prise en compte de l'information temporelle avec STM-S par rapport au M-S standard. Ensuite, nous comparons la précision du filtrage STM-S par rapport aux centroïdes trouvés avec les k-means, méthode classiquement utilisée pour le traitement de séries temporelles. Pour terminer, nous comparons les résultats des groupements obtenus par STM-S et par k-means.

7.3.1 Mesures quantitatives

La qualité du filtrage est évaluée avec un rapport signal sur bruit ($PSNR$) adapté aux données spatiotemporelles et multidimensionnelles. Cette mesure dépend de l'erreur quadratique moyenne à chaque instant entre le résultat du filtrage $\hat{\mathbf{X}}$ et les évolutions de référence $\mathbf{X}^*$:

$$PSNR(\hat{\mathbf{X}}, \mathbf{X}^*) = 10 \log_{10} \left(\frac{\mathrm{MAX}^2_{\mathbf{X}_{in}}}{\frac{1}{nT} \sum_{i=1}^{n} \|\hat{\mathbf{x}}_{\mathrm{t},i} - \mathbf{x}^*_{\mathrm{t},i}\|^2} \right) \tag{7.13}$$

avec $\mathbf{X}_{in}$ la série d'images à filtrer et $\mathrm{MAX}_{\mathbf{X}_{in}}$ la dynamique de ses mesures dans l'espace amplitude. Dans la mesure où la procédure M-S traite un seul point temporel à la fois, les résultats obtenus à chaque instant du temps ont été concaténés afin de reconstituer l'évolution filtrée des niveaux de gris. Par conséquent, le $PSNR$ est calculé de la même manière qu'il s'agisse du résultat obtenus par filtrage STM-S ou par filtrage M-S.

La classification $\mathcal{C}_{\mathbf{X}}$ obtenue par l'approche STM-S est évaluée avec le calcul du coefficient de $DICE$ pour chaque classe r par rapport à la référence $\mathcal{C}_{\mathbf{X}^*}$.

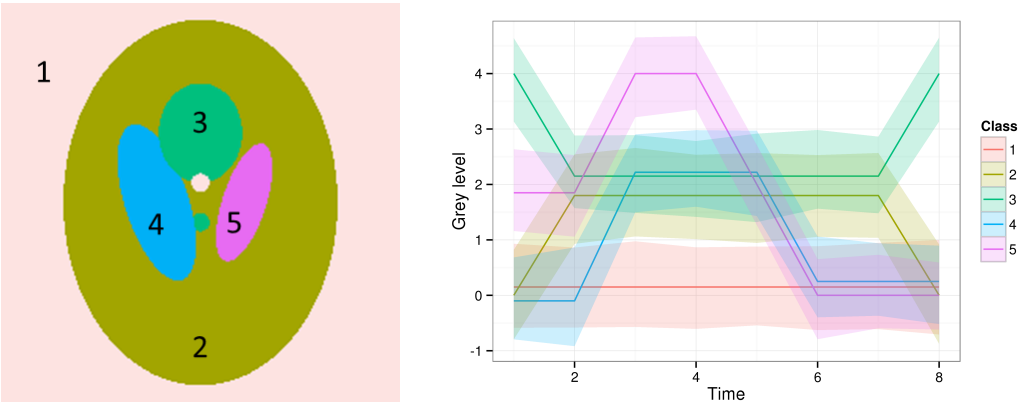
$$DICE(\mathcal{C}^r_{\mathbf{X}}, \mathcal{C}^r_{\mathbf{X}^*}) = \frac{2|\mathcal{C}^r_{\mathbf{X}} \cap \mathcal{C}^r_{\mathbf{X}^*}|}{|\mathcal{C}^r_{\mathbf{X}}| + |\mathcal{C}^r_{\mathbf{X}^*}|} \tag{7.14}$$

avec $|\cdot|$ le cardinal d'un ensemble. On rappelle que l'approche STM-S n'a pas d'a priori sur le nombre de classes.

7.3.2 Données simulées

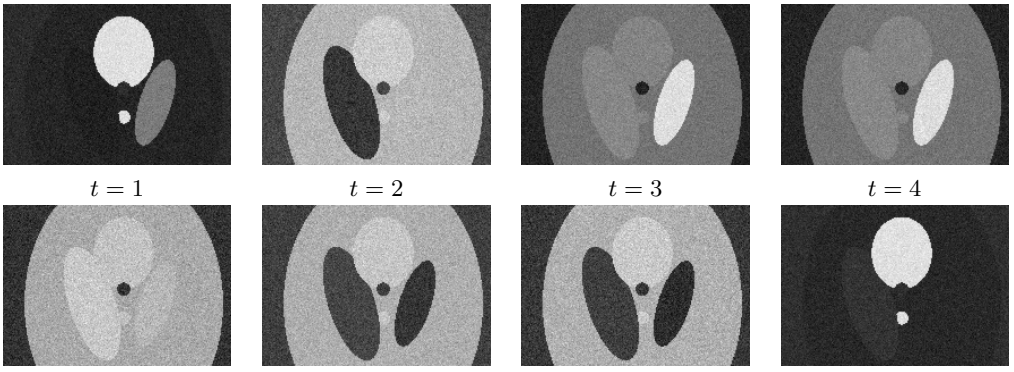
Afin de tester l'efficacité de l'approche STM-S, nous avons simulé des séquences d'images à la fois en niveaux de gris et en couleur. Le premier jeu de données est une séquence d'images de taille 256×256 pixels basée sur le fantôme de Shepp-Logan [58]. Nous appellerons désormais ce jeu de données la séquence Shepp-Logan (S-L). A chacune des cinq classes de la vérité terrain, présentées Fig.7.5(a), a été associé un modèle d'évolution (Fig.7.5(b)) afin de générer un jeu de données simulé de taille $256 \times 256 \times 8$ dont la troisième dimension est définie comme l'espace temps. Ensuite chaque image, correspondant à un point temporel différent, a été dégradée via une convolution avec un filtre gaussien de taille 10×10 et d'écart-type 0.4 puis par l'addition d'un bruit gaussien centré et d'écart-type 0.2 (Fig.7.6). Ces dégradations modifient suffisamment les niveaux de gris pour que des intensités de pixels appartenant à des classes différentes soient mélangées à certains points temporels (Fig.7.5(b)). De plus, les évolutions ont été simulées de telle sorte que les classes ne puissent pas être retrouvées sans utiliser toute l'information temporelle.

Nous avons aussi simulé cinq classes d'évolution d'intensités rouge, vert et bleu des pixels d'une image couleur (Fig.7.7(a)), que nous appellerons séquence image couleurs (IC). Une série d'images 64×64 de 8 instants a été générée. Il est important de préciser que les évolutions choisies pour les intensités ne permettent pas la séparation des cinq classes sans prendre en compte simultanément les évolutions d'intensités dans les trois canaux couleur. Ces images ont ensuite été dégradées à l'aide d'un filtre moyen de taille 3×3 puis par l'addition d'un bruit gaussien centré et d'écart-type 0.3, appliqués indépendamment sur chaque dimension couleur et à chaque instant. La figure 7.7(b) illustre l'évolution de la séquence IC. De plus, les dégradations introduites rendent difficile la détermination des frontières des classes.



(a) (b)

FIGURE 7.5 – (a) Vérité terrain de la séquence Shepp-logan. (b) Evolutions simulées. Les couleurs correspondent aux classes en (a). Les couleurs transparentes autour des modèles d'évolution représentent la variabilité introduite par la dégradation des images.



$t = 1$ $t = 2$ $t = 3$ $t = 4$

$t = 5$ $t = 6$ $t = 7$ $t = 8$

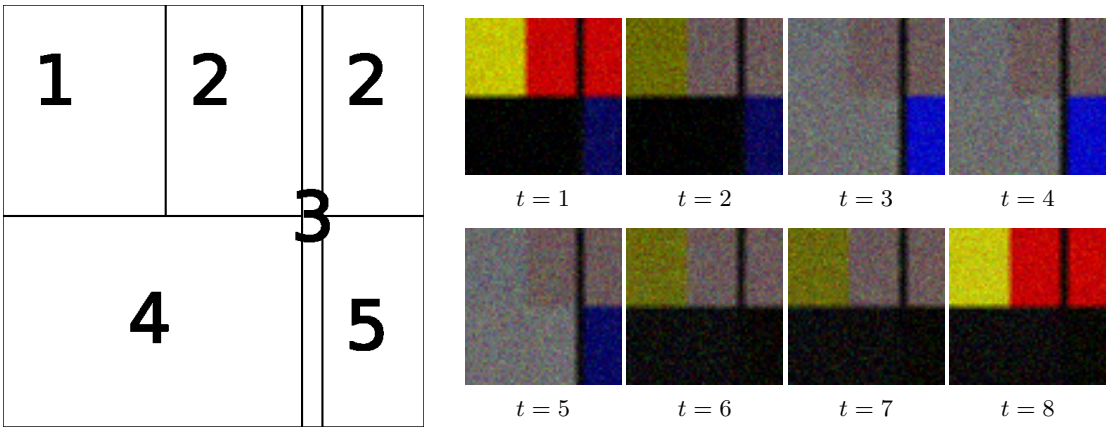
FIGURE 7.6 – Séquence Shepp-logan créée à partir des classes de la vérité terrain (Fig.7.5(a)) et des modèles d'évolutions (Fig.7.5(b)) dégradés par un filtre gaussien et un bruit additif de type gaussien.

7.3.3 Evaluation de la qualité du filtrage

Dans un premier temps, nous évaluons le bénéfice apporté par la prise en compte de l'information temporelle avec STMS par rapport au formalisme M-S classique. Ensuite, nous présentons l'intérêt de STM-S par rapport à l'approche k-means étendue aux séries temporelle, qui est un algorithme de référence pour le groupement de données temporelles.

STM-S vs. MS

La figure 7.8 expose les résultats obtenus sur la séquence S-L par les procédures M-S et STM-S après sélection, respectivement, de leurs paramètres d'échelles optimaux. On peut observer sur le résultat du filtrage obtenu avec l'approche M-S que les niveaux de gris de certaines classes sont mélangés (flèches bleues dans Fig.7.8), des contours sont morcelés entre deux régions quand leurs niveaux de gris sont suffisamment proches pour être mé-



(a) (b)

FIGURE 7.7 – (a) Vérité terrain. (b) Séquence IC générée après dégradations par le filtre moyenneur et l'ajout de bruit gaussien.

langés par les dégradations introduites dans l'image (flèches rouges dans Fig.7.8) où des points aberrants restent non filtrés (flèches vertes dans Fig.7.8). Néanmoins, l'évolution moyenne à l'intérieur de chaque classe correspond avec le modèle d'évolution qui lui est associé dans la vérité terrain (Fig.7.9). Concernant le résultat du filtrage obtenu par l'approche STM-S, les évolutions moyennes à l'intérieur des classes correspondent aussi aux modèles de la vérité terrain tout en ayant une variabilité intra-classe quasi nulle contrairement à l'approche M-S (Fig.7.9). De plus, les régions associées aux classes sont bien délimitées avec l'approche STM-S même pour celles dont les niveaux de gris sont proches. Les contours sont préservés et on peut noter l'absence de points aberrants. Bien que chacune des procédures améliore nettement la qualité globale de la séquence, en utilisant les paramètres d'échelles optimaux, l'approche STM-S dépasse de loin l'approche M-S d'une part en qualité de filtrage et d'autre part en similarité avec les modèles d'évolutions de la vérité terrain. Il est évident, après comparaison des figures 7.5(b) et 7.9, que le filtrage STM-S permet d'obtenir les évolutions de niveaux de gris les plus facilement distinguables et les mieux séparées. Les valeurs maximales du $PSNR$ sont obtenues avec des paramètres d'échelle dans l'espace amplitude variant entre 0,4 et 1,0 (Fig. 7.10). Dans cet intervalle de valeurs, l'approche STM-S permet d'obtenir de plus hautes valeurs de $PSNR$ que l'approche M-S pour la plupart des échelles spatiales. Les valeurs optimales de $PSNR$ obtenues respectivement pour STM-S sont 62dB ($h_\mathrm{s} = 255$, $h_\mathrm{r} = 0.4$) et pour M-S de 45dB ($h_\mathrm{s} = 3$, $h_\mathrm{r} = 0.5$). Par conséquent, la prise en compte de l'information temporelle avec STM-S (7.4) permet d'augmenter le $PSNR$ de 21dB.

De plus, même si l'échelle en amplitude n'est pas choisie de manière optimale, la qualité du filtrage STM-S descend rarement en-dessous de celle du filtrage M-S pour des paramètres d'échelles identiques. La contrainte très stricte de similarité temporelle introduite par la norme infini (7.6) permet de paramétrer la méthode avec une grande échelle spatiale, ce qui n'est pas le cas avec M-S car chaque point temporel est considéré indépendamment

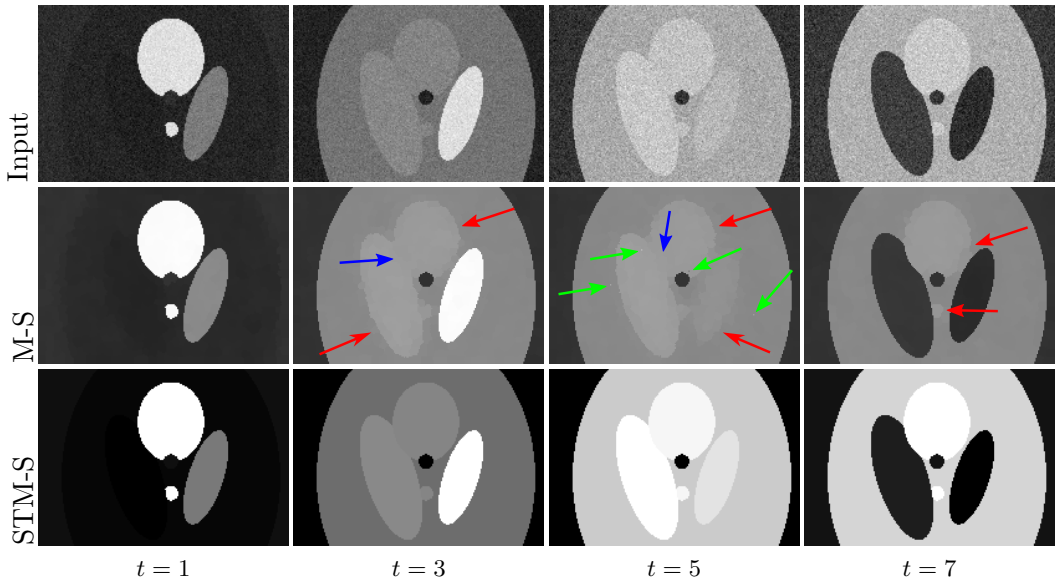


FIGURE 7.8 – Première ligne : quatre images de la séquence S-L. Seconde ligne : résultats obtenus par filtrage M-S avec sa combinaison d'échelles optimale ($h_s = 3$ pixels, $h_r = 0.5$) aux temps 1, 3, 5, 7. Troisième ligne : résultats obtenus par filtrage STM-S avec sa combinaison d'échelles optimale ($h_s = 255$ pixels, $h_r = 0.4$) pour les mêmes points temporels. Les flèches rouges pointent les contours flous. Les flèches bleues indiquent les classes mélangées. Les flèches vertes indiquent les points aberrants restants.

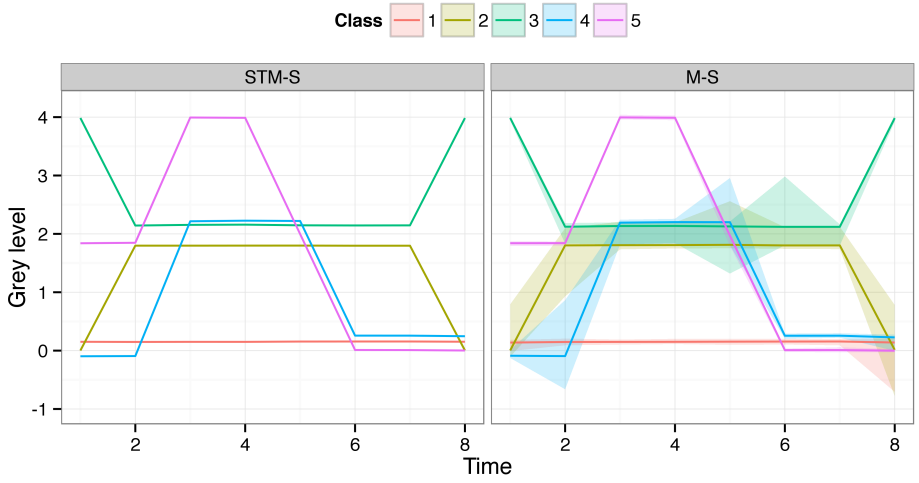
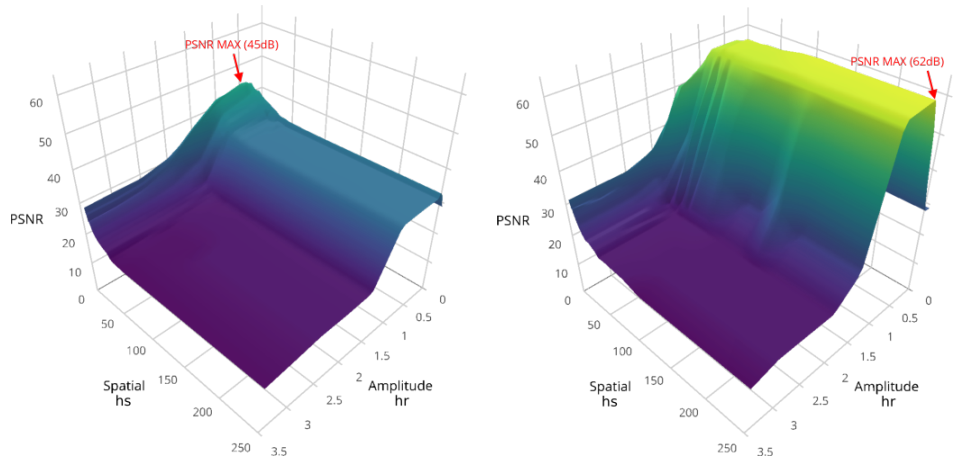


FIGURE 7.9 – Evolutions des intensités après filtrage STM-S (gauche) et M-S (droite). Les lignes en couleur représentent l'évolution moyenne de chaque classe. Les couleurs transparentes représentent la variablilité des évolutions dans chaque classe.

des autres, *i.e.* aucune discrimination sur la similarité des évolutions dans l'espace amplitude n'est possible. Il en résulte que les paramètres optimaux pour l'approche M-S sont de petits paramètres d'échelles dans les domaines spatial et amplitude tandis que les paramètres optimaux pour l'approche STM-S sont une petite échelle en amplitude combinée à une grande échelle spatiale. C'est le fait que STM-S puisse prendre en compte

plus d'échantillons avec des évolutions similaires dans l'espace amplitude grâce à une plus grande échelle spatiale qui permet d'améliorer le $PSNR$.



(a) (b)

FIGURE 7.10 – $PSNR$ tracé en fonction des combinaisons d'échelles dans les domaines spatial et amplitude pour M-S (gauche) et STM-S (droite) sur les donnée Shepp-Logan.

Nous présentons maintenant les résultats obtenus avec la séquence IC par l'extension du formalisme STM-S aux mesures multidimensionnelles (7.9) et(7.10). La séquence d'images RGB est présentée avant et après filtrage STM-S figure 7.11, avec des matrices échelles $\mathbf{H}_\mathrm{s} = diag(\infty, \infty)$ et $\mathbf{H}_\mathrm{r} = diag(1.75^2, 1.75^2, 1.75^2)$. L'utilisation de très grandes valeurs pour l'échelle spatiale permet de regrouper les échantillons en tenant uniquement compte de leurs évolutions dans l'espace amplitude sans introduire de contrainte sur le voisinage spatial. La méthode exploite ainsi toutes les évolutions dans l'espace amplitude pour filtrer chacun des échantillons. On peut remarquer que malgré les dégradations introduites, le résultat du filtrage STM-S est visuellement satisfaisant (Fig.7.11). Quantitativement, le $PSNR$ est égal à 19dB pour la séquence d'entrée est de 31dB après filtrage.

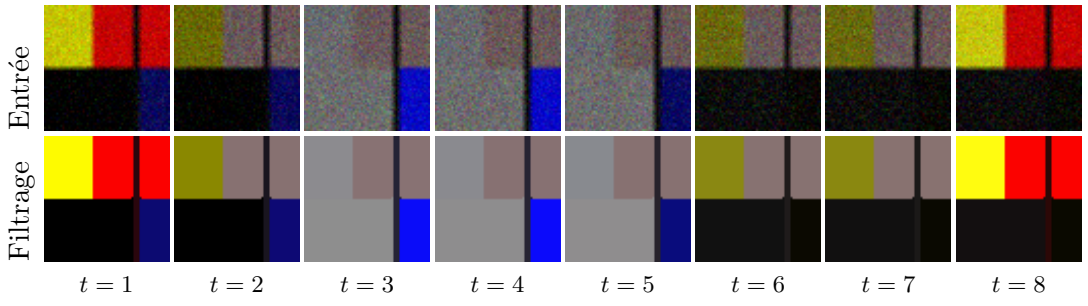


FIGURE 7.11 – Première ligne : séquence d'images RGB simulées et bruitées. Deuxième ligne : résultats du filtrage STM-S avec $\mathbf{H}_\mathrm{s} = diag(\infty, \infty)$ et $\mathbf{H}_\mathrm{r} = diag(1.75^2, 1.75^2, 1.75^2)$.

STM-S vs. k-means

Nous avons montré que la prise en compte de l'information temporelle dans le processus de filtrage permet d'améliorer considérablement la précision de l'approche M-S. Nous voulons maintenant déterminer si la méthode proposée est plus efficace que les k-means, qui contrairement au M-S utilise la distance euclidienne entre deux séries temporelles $\mathbf{u}_t$ et $\mathbf{v}_t$:

$$Sim(\mathbf{u}_t, \mathbf{v}_t) = L_2(\mathbf{u}_t, \mathbf{v}_t) \tag{7.15}$$

Afin de tester la qualité du filtrage par k-means nous avons tiré aléatoirement 1000 combinaisons de 5 centres de classes dans la séquence Shepp-Logan. Nous pouvons observer figure 7.12 la distribution des $PSNR$ après filtrage k-means sur les données brutes. La

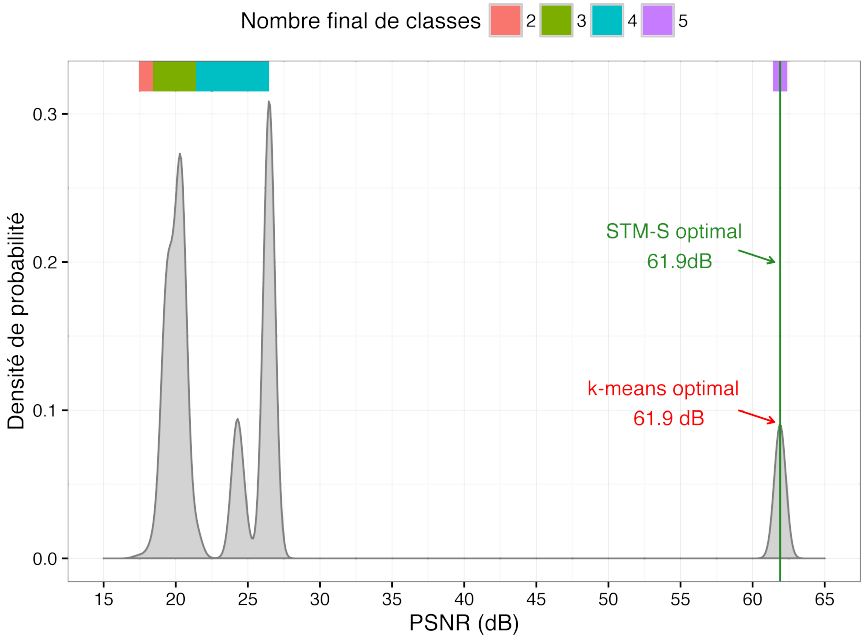


FIGURE 7.12 – Densité de probabilité du PSNR après filtrage k-means sur les données simulées pour 1000 initialisations aléatoires de 5 centroïdes. Barre couleur : nombre de classes non vides après filtrage k-means.

prise en compte de l'information temporelle permet aux k-means d'atteindre une qualité de filtrage supérieure au M-S et égale à STM-S dans les meilleurs cas. D'après la barre couleur au-dessus de la densité de probabilité nous pouvons voir que les k-means, bien que toujours initialisés à 5 classes, convergent dans 90% des cas vers une solution où certains centroïdes finissent dans un ensemble vide. La valeur du $PSNR$ est d'ailleurs corrélée au nombre de centroïdes, plus la qualité du filtrage est faible et plus le nombre de centroïdes non affectés est important.

Nous observons aussi que les résultats optimaux ne sont obtenus que pour moins de 10% des initialisations. De plus, la zone d'intérêt traitée dans cet exemple est de 202×142 pixels soit 28684 échantillons. Ce problème est sous dimensionné par rapport à un cas réel avec possiblement beaucoup plus de 5 classes à retrouver. Par conséquent, le faible taux

de réussite de l'approche k-means illustre les difficultés d'utilisation et de reproductibilité de cette approche par rapport à STM-S.

Effectivement, il suffit d'estimer la distribution des caractéristiques afin de bien paramétrer le critère de similarité temporelle h_{r} de STM-S, celui qui a le plus fort impact sur la qualité résultat final (Fig. 7.10). La valeur du paramètre d'échelle spatiale h_{s} permet quant à lui de déterminer si l'étude est faite localement ou plus globalement, mais n'influe que très peu sur la qualité finale du filtrage.

Pour terminer, les résultats de groupement après filtrage que ce soit avec STM-S ou k-means sont parfaits quand les valeurs de $PSNR$ sont optimales. Si on se réfère à la densité de probabilité du résultat en Fig. 7.12 cela correspond à seulement 10% des fois où k-means a été relancé, ce qui laisse 90% de chance d'avoir un résultat non satisfaisant. Rappelons aussi que la connaissance du nombre de classes est nécessaire pour utiliser k-means et que la qualité du résultat dépend fortement de l'initialisation, ce n'est pas un algorithme déterministe. De plus, il n'est généralement pas envisageable de déterminer les centroïdes "à la main" en pratique sur des données réelles.

TABLE 7.1 – Coefficients DICE des 5 classes obtenues par k-means et STM-S sur les données simulées. Première colonne : coefficients de $DICE$ obtenus pour une valeur non optimale du $PSNR$ avec la méthode k-means. Colonnes trois et quatre : coefficients de $DICE$ obtenus avec les $PSNR$ optimaux de chaque méthode.

Classe	k-means 19.4dB (27% des cas)	k-means optimal (10% des cas)	STM-S optimal
1	1.00	1.00	1.00
2	1.00	1.00	1.00
3	-	1.00	1.00
4	0.42	1.00	1.00
5	-	1.00	1.00

7.4 Conclusion

Dans ce chapitre, nous avons proposé une nouvelle approche appelée STM-S étendant le formalisme mean-shift au filtrage et au groupement de séries temporelles et démontré ses apports au M-S standard. La prise en compte de l'information temporelle permet de retrouver des évolutions qui se rapprochent au maximum des signatures temporelles originales des pixels. Les valeurs optimales du $PSNR$ obtenues avec STM-S sont bien plus hautes que celles obtenues avec M-S sur la séquence d'images simulées Shepp-Logan. Deuxièmement, nous avons montré que la méthode STM-S est une réelle alternative à la méthode de groupement standard de séries temporelles k-means. Notre approche présente plusieurs avantages : (1) Pas de connaissance nécessaire a priori sur le nombre de classes à retrouver ni sur la structure des classes à retrouver. (2) Paramétrage intuitif de deux valeurs : les critères de similarité spatiale et temporelle. (3) Méthode déterministe après le choix des paramètres d'échelle. Nous avons montré sur la séquence d'images simulées S-L que STM-S permet d'obtenir, dans 90% des cas, un filtrage (valeurs de $PSNR$) et

des groupes (coefficients de $DICE$) de séries temporelles plus satisfaisants qu'avec une approche classique de type k-means.

Les comparaisons entre séries temporelles effectuées par STM-S étant basées sur la distance euclidienne, cette approche n'est pas robuste aux distorsions temporelles qu'il pourrait exister entre elles. C'est le verrou que nous essayons de lever dans le prochain chapitre.

Chapitre 8

Prise en compte des déphasages temporels : utilisation des appariements obtenus par la DTW

8.1 Introduction

Dans le chapitre précédent, nous avons présenté une méthode permettant l'identification de groupes de pixels dont les caractéristiques évoluent similairement et de manière synchrone au cours du temps. Cependant, deux évolutions identiques peuvent être déphasées ou présenter des distorsions temporelles locales. Dans la mesure où le critère de similarité de STM-S est basé sur l'écart maximal calculé temps à temps entre deux courbes, deux évolutions identiques mais n'étant pas synchronisées seront associées à deux groupes différents.

Nous proposons dans ce chapitre un algorithme de groupement hiérarchique avec une contrainte spécifique sur les appariements optimaux déterminés par le calcul de la DTW. L'apport de cet algorithme est de grouper les séries temporelles déphasées tout en ajoutant une contrainte sur les écarts locaux des appariements trouvés avec la DTW (Sakoe [10,11]).

8.2 Algorithme de groupement

Pour rappel la DTW, présentée en section 2.5.2, est formulée comme ceci :

$$DTW(\mathbf{u},\mathbf{v}) = \min_{\mathbf{p}_l} \left(\frac{\sum_{l=1}^{\mathrm{L}} d(\mathbf{u},\mathbf{v},\mathbf{p}_l)\cdot w_l}{\sum_{l=1}^{\mathrm{L}} w_l} \right) = \frac{\sum_{l=1}^{\mathrm{L}} d(\mathbf{u},\mathbf{v},\mathbf{p}_l^*)\cdot w_l}{\sum_{l=1}^{\mathrm{L}} w_l} \tag{8.1}$$

avec w_l les poids associés à chaque appariement, $\mathbf{p}_l^*$ les appariements optimaux entre deux séries temporelles ou $\mathbf{p}_l$ les appariements entre deux séries temporelles et d la distance euclidienne entre deux éléments associés :

$$d(\mathbf{u},\mathbf{v},\mathbf{p}_l) = \| \mathbf{p}_l(\mathbf{u}) - \mathbf{p}_l(\mathbf{v}) \| \tag{8.2}$$

Le chemin d'alignements optimal P^* est déterminé de la manière suivante :

$$P^*_{\mathbf{u},\mathbf{v}} = arg\min_{P} \left(\frac{\sum_{l=1}^{\mathrm{L}} d(\mathbf{u},\mathbf{v},\mathbf{p}_l)\cdot w_l}{\sum_{l=1}^{\mathrm{L}} w_l} \right) \tag{8.3}$$

Les versions alignées de $\mathbf{u}$ et $\mathbf{v}$, par rapport à P^*, sont notées respectivement $P^*_{\mathbf{u},\mathbf{v}}(\mathbf{u})$ et $P^*_{\mathbf{u},\mathbf{v}}(\mathbf{v})$. Les notations raccourcies de ces deux vecteurs sont $P^*(\mathbf{u})$ et $P^*(\mathbf{v})$. P^* est quand à lui composé de L couples appariés, $\mathbf{p}_l^*(\mathbf{u})$ et $\mathbf{p}_l^*(\mathbf{v})$ correspondant aux $l^{ièmes}$ mesures des séquences $\mathbf{u}$ et $\mathbf{v}$ une fois alignées.

Comme formulé en (8.1), la DTW intègre les écarts entre les valeurs appariées de chaque série temporelle. Par conséquent, la valeur de la DTW ne permet d'avoir qu'une idée de la dissimilarité globale entre deux séries temporelles sans garantir de dissimilarité locale entre les valeurs appariées. Avant de grouper les échantillons notre but est de déterminer si localement, *i.e.* pour chacun des appariements optimaux, des évolutions de mesures dans le domaine amplitude sont proches. Pour ce faire, nous appliquons la même contrainte sur l'écart entre les appariements des séries temporelles qu'en (7.6) avec à la norme infini. La différence est qu'elle n'est pas calculée sur les écarts temps à temps entre deux évolutions mais sur les écarts entre les appariements optimaux P^* déterminés par la DTW. Cette nouvelle contrainte notée d_∞ est définie de la manière suivante :

$$d_\infty(\mathbf{u},\mathbf{v},P^*) = \| \mathbf{H}_{\mathrm{r}}^{-\frac{1}{2}} \left(P^*(\mathbf{u}) - P^*(\mathbf{v}) \right) \|_\infty \tag{8.4}$$

Cette mesure permet de quantifier l'écart maximal qu'il existe entre deux séries temporelles appariées. L'algorithme de groupement, cf. algorithme 5, utilise la fonction de pondération suivante :

$$g\Big(d_\infty^2(\cdot)\Big) = \begin{cases} 1 & \text{si } d_\infty^2(\cdot) \leq 1 \\ 0 & \text{sinon} \end{cases} \tag{8.5}$$

La fonction de pondération $g(\cdot)$ peut être interprétée comme un critère de décision pour la question : deux séries temporelles sont-elles proches ? Si l'écart maximal en amplitude entre les appariements optimaux trouvés par la DTW ne dépasse pas le seuil fixé par $\mathbf{H}_{\mathrm{r}}$, alors les deux séries temporelles sont considérées similaires. L'algorithme permettant de grouper les échantillons est présenté ci-dessous :

Algorithm 5 Partitionnement des échantillons en utilisant les appariements optimaux trouvés avec la DTW entre deux séries temporelles

Require: $\hat{\mathbf{X}} = \{\hat{\mathbf{x}}_c\}_{c=1\ldots C}$: Evolutions simulées
$\hat{\mathbf{Q}} = \{\hat{q}_c\}$: Initialisation des classes comme singletons
repeat
  $permut \leftarrow$ false
  for all $\hat{\mathbf{x}}_c \in \hat{\mathbf{X}}$ **do**
    $(\hat{\mathbf{x}}_j^*, P^*) = arg\min_{\hat{\mathbf{x}}_j \in \hat{\mathbf{X}} \setminus \hat{\mathbf{x}}_c} D(\hat{\mathbf{x}}_{\mathrm{t},c}, \hat{\mathbf{x}}_{\mathrm{t},j}, P^*)$
    $\mathcal{N} \leftarrow \{\, k \in [1, C] \mid \hat{q}_k = \hat{q}_j \,,\, k \neq c \,\}$
    $isNeighbor \leftarrow$ true
    for all $k \in \mathcal{N}$ **do**
      if $g\Big(d_\infty\left(\hat{\mathbf{x}}_{\mathrm{t},c}, \hat{\mathbf{x}}_{\mathrm{t},k}^*, P^*\right)\Big) \neq 1$ **then**
        $isNeighbor \leftarrow$ false
      end if
    end for
    if $isNeighbor$ is true **then**
      $\hat{q}_c \leftarrow \hat{q}_j$
      $permut \leftarrow$ true
    end if
  end for
until $permut$ is false
$\mathbf{Q}^* \leftarrow \hat{\mathbf{Q}}$
return $\mathbf{Q}^*$

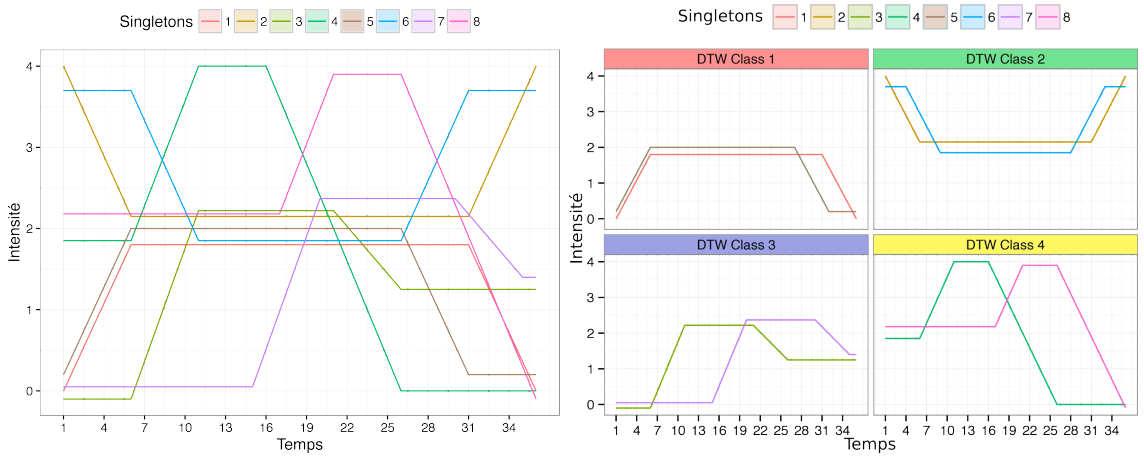
Dans cet algorithme des échantillons sont groupés s'ils sont proches au regard de (8.5) et si l'échantillon candidat est aussi proche de tous les autres échantillons du groupe auquel est associée la référence. Ceci permet de limiter "l'effet de lien" pendant la phase de groupement. Par exemple si deux classes se recouvrent partiellement, cette technique évite de les fusionner à cause d'associations des échantillons de proche en proche. Ce critère peut être interprété comme un moyen de limiter la variance inra-classe des groupes. Néanmoins, cet algorithme fonctionnera bien si les groupes sont compacts mais risque d'être sensible aux échantillons aberrants ou s'éloignant trop des centres des groupes.

8.3 Validation : données simulées et résultats

Dans cette section, l'algorithme de groupement est testé sur des séries présentant des distorsions sur l'axe temporel. Ce jeu de données simulées a été créé afin d'évaluer la capacité de la méthode à grouper des évolutions déphasées et légèrement déformées. Il est constitué de huit séries temporelles considérées comme des singletons à l'entrée de l'algo-

rithme de groupement. La figure 8.1(a) montre les séries temporelles générées, elle devront être séparée en 4 groupes de deux série temporelles. Néanmoins, les évolutions censées correspondre après l'algorithme de groupement ne sont pas identiques. Leurs signatures temporelles sont légèrement dissimilaires en amplitude, déphasées et/ou étirées. Nous nous attendons à ce que l'algorithme trouve sans a priori le nombre de groupe et les bonnes correspondances entre les séries temporelles.

On peut voir en Fig. 8.1(b) les associations trouvées par l'algorithme. Le critère utilisé sur les associations de la DTW permet de retrouver des signatures temporelles similaires bien que désynchronisées : les séries temporelles sont correctement mises en correspondance et le bon nombre de groupes est retrouvé sans avoir besoin de le spécifier préalablement.



(a) (b)

FIGURE 8.1 – (a) Séries temporelles simulées. (b) Groupements obtenus avec l'algorithme proposé. Le critère de similarité a été fixé à $\mathbf{h}_r = 0.4$.

8.4 Conclusion

Nous avons présenté un algorithme qui permet de grouper des séries temporelles, invariant à des distorsions sur l'axe temporel. Son originalité est d'utiliser une contrainte nouvelle sur les appariements optimaux obtenus avec la mesure DTW. L'utilisation de la norme infini, comme en section 7.2.1, permet de contraindre l'écart maximal toléré entre deux associations. Ce critère de groupement garanti une similarité locale, au seuil fixé près, entre les associations de la DTW. De plus, elle permet de ne pas avoir a spécifier le nombre de classes à retrouver. En effet, tant que des séries temporelles similaires sont identifiées elles vont être associées au même groupe, une fois que plus aucune série n'est voisine à un groupe alors le processus de fusion s'arrêtera de lui-même. Nous avons démontré le bon fonctionnement de notre algorithme sur un jeu de données simulées, sans avoir spécifié au préalable le nombre de groupes à retrouver ni de critère d'arrêt pour l'association des séries temporelles.

Jusqu'alors la représentation des données des algorithmes proposés prend en compte l'ordonnancement temporel des mesures mais ne prend pas explicitement en compte les valeurs du temps leur étant associées. En effet, seul l'ordre des mesures dans les vecteurs $\mathbf{x}_{\mathrm{t},i}$ et la fréquence d'échantillonnage entre séries temporelles étaient importants à respecter. Cependant, ceci peut être un point limitant lorsque l'instant d'acquisition peut apporter des informations précieuses comme dans le cas où la vitesse d'évolution des mesures pourrait être un critère discriminant. C'est pourquoi nous proposons dans le chapitre suivant une extension du formalisme mean-shift prenant explicitement en compte la valeur du temps pour le filtrage des échantillons.

Chapitre 9

Prise en compte des déphasages temporels : STM-S^{++}

9.1 Introduction

Dans les chapitres précédents, nous avons présenté deux algorithmes. L'un permettant d'identifier des échantillons dont les caractéristiques évoluent de manière similaire et synchronisée et l'autre capable de grouper les échantillons dont les évolutions peuvent être déphasées et/ou déformées sur l'axe temporel. La principale limitation de ces deux premières contributions est l'hypothèse que la fréquence d'échantillonnage entre les séries temporelles est considérée identique. En effet, les valeurs temporelles ne sont pas utilisées dans les formalismes présentés. Ainsi, la notion de chronologie n'entre pas en compte avec STM-S, il suffit que les échantillons correspondants soient stockés au même indice dans le vecteur temporel. Concernant l'algorithme de groupement présenté dans le chapitre précédent seule la notion d'ordre importe, les valeurs des instants temporels sont supposés identiques et les séries temporelles ordonnées en temps. Ceci peut constituer un verrou important lorsqu'il s'agit de comparer des séries temporelles acquises via des systèmes et des protocoles différents, cas fréquents avec les applications médicales qui nous préoccupent.

Dans ce chapitre, nous présentons une extension du formalisme M-S prenant explicitement en compte la valeur temporelle des échantillons ce qui permet de lever ces limitations. Après avoir exposé le formalisme de la méthode nous la testons sur des séquences d'images réelles et montrons son intérêt dans le cas d'occultation d'objets.

9.2 Méthode

Dans cette section, nous proposons un formalisme unifiant mean-shift causal et omniscient introduits section 4.2.2, appelé STM-S^{++}, grâce à deux paramètres d'échelle permettant d'ajuster la fenêtre temporelle à prendre en compte.

Considérons un ensemble de n échantillons aux positions $\{\mathbf{x}_{\mathrm{s},i}\}_{i=1\ldots n}$, un ensemble de caractéristiques dans le domaine amplitude $\{\mathbf{x}_{\mathrm{r},i}\}_{i=1\ldots n}$ ainsi qu'un ensemble de valeurs scalaires $\{x_{\mathrm{t},i}\}_{i=1\ldots n}$ représentant le temps. Notons les tailles des supports spatial et amplitude $\mathcal{S}$ et $\mathcal{R}$, respectivement. L'ensemble d'entrée $\mathbf{X} = \{\mathbf{x}_i\}_{i=1\ldots n}$ est par conséquent défini comme :

$$\mathbf{x}_i = \left[\mathbf{x}_{\mathrm{s},i}^{'}\ \mathbf{x}_{\mathrm{r},i}^{'}\ x_{\mathrm{t,\,i}}\right]^{'} \qquad \text{avec} \qquad \begin{array}{l} \mathbf{x}_{\mathrm{s},i} \in \mathbb{R}^{\mathcal{S}}, \mathbf{x}_{\mathrm{r},i} \in \mathbb{R}^{\mathcal{R}}, x_{\mathrm{t},i} \in \mathbb{R} \\ i = 1, \ldots, n : \text{ échantillons} \end{array} \tag{9.1}$$

Dans ce nouveau formalisme, la valeur du temps est explicitement prise en compte avec x_t, ce qui n'était pas le cas avec les méthodes proposées précédemment. Basé sur ces notations, nous proposons l'équation suivante pour calculer itérativement le filtrage mean-shift spatio-temporel des échantillons $\mathbf{x}_i^{[\mathrm{k}+1]}$:

$$\mathbf{x}_i^{[\mathrm{k}+1]} = \frac{\sum_{j=1}^{n} S_{i,j}(\cdot)\,.R_{i,j}(\cdot)\,.T_{i,j}(\cdot)\,.\mathbf{x}_j^{[\mathrm{k}]}}{\sum_{j=1}^{n} S_{i,j}(\cdot)\,.R_{i,j}(\cdot)\,.T_{i,j}(\cdot)} \tag{9.2}$$

où $S_{i,j}(\cdot)$, $R_{i,j}(\cdot)$ et $T_{i,j}(\cdot)$ sont respectivement les distances spatiale, en amplitude et temporelle pondérées entre un échantillon référence $\mathbf{x}_i$ et un candidat $\mathbf{x}_j$ ($\mathbf{x}_i$ et $\mathbf{x}_j \in \mathbb{R}^{\mathcal{S}+\mathcal{R}+1}$) :

$$S_{i,j}\Big(\mathbf{x}_{\mathrm{s},i}^{[\mathrm{k}]}, \mathbf{x}_{\mathrm{s},j}^{[\mathrm{k}]}, \mathbf{H}_{\mathrm{s}}\Big) = g_{\mathrm{s}}\Big(\, d_{\mathrm{s}}^2\Big(\mathbf{x}_{\mathrm{s},i}^{[\mathrm{k}]}, \mathbf{x}_{\mathrm{s},j}^{[\mathrm{k}]}, \mathbf{H}_{\mathrm{s}}\Big)\Big) \tag{9.3}$$

$$R_{i,j}\Big(\mathbf{x}_{\mathrm{t},i}^{[\mathrm{k}]}, \mathbf{x}_{\mathrm{t},j}^{[\mathrm{k}]}, \mathbf{H}_{\mathrm{r}}\Big) = g_{\mathrm{r}}\Big(\, d_{\mathrm{r}}^2\Big(\mathbf{x}_{\mathrm{t},i}^{[\mathrm{k}]}, \mathbf{x}_{\mathrm{t},j}^{[\mathrm{k}]}, \mathbf{H}_{\mathrm{r}}\Big)\Big) \tag{9.4}$$

$$T_{i,j}\Big(x_{\mathrm{t},i}^{[\mathrm{k}]}, x_{\mathrm{t},j}^{[\mathrm{k}]}, h_{\mathrm{t}^-}, h_{\mathrm{t}^+}\Big) = G_{\mathrm{t}}\Big(\, \varepsilon_{\mathrm{t}}\Big(x_{\mathrm{t},i}^{[\mathrm{k}]}, x_{\mathrm{t},j}^{[\mathrm{k}]}\Big), h_{\mathrm{t}^-}, h_{\mathrm{t}^+}\Big) \tag{9.5}$$

Dans ces équations, $d_{\mathrm{s}}(\mathbf{u}_{\mathrm{s}}, \mathbf{v}_{\mathrm{s}}, \mathbf{H}_{\mathrm{s}})$ et $d_{\mathrm{r}}(\mathbf{u}_{\mathrm{r}}, \mathbf{v}_{\mathrm{r}}, \mathbf{H}_{\mathrm{r}})$ sont les distances euclidienne généralisées calculées sur les supports spatial et amplitude de deux échantillons $\mathbf{u}$ and $\mathbf{v}$. $\mathbf{H}_{\mathrm{s}}$ and $\mathbf{H}_{\mathrm{r}}$ sont les matrices échelles spatial et amplitude de dimensions $\mathcal{S}\times\mathcal{S}$ et $\mathcal{R}\times\mathcal{R}$. La distance de Mahalanobis est définie comme :

$$d(\mathbf{u}, \mathbf{v}, \mathbf{H}) = \Big((\mathbf{u}-\mathbf{v})^{'}\,\mathbf{H}^{-1}\,(\mathbf{u}-\mathbf{v})\Big)^{1/2} \tag{9.6}$$

avec $\mathbf{H}$ la matrice échelle, carrée et définie positive.

Dans le domaine temporel, les écarts entre deux échantillons $\mathbf{u}$ et $\mathbf{v}$ sont calculés afin

de prendre en considération l'ordre temporel :

$$\varepsilon_{\mathrm{t}}(u_{\mathrm{t}}, v_{\mathrm{t}}) = (v_{\mathrm{t}} - u_{\mathrm{t}}) \tag{9.7}$$

Ensuite, les poids associés aux échantillons dans (9.2) seront égaux à l'association des distances pondérées sur les supports spatial, amplitude et temporel. Les mêmes noyaux g sont utilisés pour pondérer les distances spatiale et amplitude :

$$g_{\mathrm{s}}\Big(d_{\mathrm{s}}^2\left(\cdot\right)\Big) = g_{\mathrm{r}}\Big(d_{\mathrm{r}}^2\left(\cdot\right)\Big) = \begin{cases} 1 & \text{if } d_{\mathrm{s}}^2\left(\cdot\right),\ d_{\mathrm{r}}^2\left(\cdot\right) \leq 1 \\ 0 & \text{sinon} \end{cases} \tag{9.8}$$

En revanche la fonction de pondération temporelle est définie comme :

$$G_{\mathrm{t}}\Big(\varepsilon_{\mathrm{t}}\left(\cdot\right)\Big) = \begin{cases} 1 & \text{if } h_{\mathrm{t}^-} \leq \varepsilon_{\mathrm{t}}\left(\cdot\right) \leq h_{\mathrm{t}^+} \\ 0 & \text{sinon} \end{cases} \tag{9.9}$$

Ainsi, paramétrer $h_{\mathrm{t}^+} = 0$ permet de procéder à un filtrage causal tandis que $\{h_{\mathrm{t}^-} = -\infty, h_{\mathrm{t}^+} = +\infty\}$ permet de procéder à un filtrage omniscient des données. Toutefois, h_{t^-} et h_{t^+} peuvent être paramétrés différemment. Leurs valeurs peuvent être choisies de manière indépendante pour ajuster les quantités d'information passée et future à prendre en compte pour le filtrage. Le principe de convergence de STM-S^{++} est illustré Fig. 9.1.

9.3 Validation

Nous venons de présenter un formalisme mean-shift unifiant les approches causale et omnisciente. Dans cette section nous présentons les résultats obtenus sur deux jeux de données réelles choisis pour mettre en avant la robustesse de la méthode aux occultations totales d'objets. L'algorithme de groupement utilisé après filtrage est le même que pour l'approche STM-S (section 7.2.3).

9.3.1 Séquence tennis de table

Nous avons choisi la vidéo *tennis de table* disponible sur internet et déjà utilisée par Gu *et al.* [50]. L'efficacité de la procédure STM-S^{++} a été testée sur un extrait de cette séquence. Une occultation a été simulée sur chacune d'elles afin de masquer complètement la balle à un instant donné (Fig.9.2).

Comme suggéré par Li *et al.* [59], les caractéristiques RGB des images ont été transposées dans l'espace Lab avant le filtrage. Les paramètres d'échelles spatiale et amplitude ($\mathbf{H}_{\mathrm{s}}$, $\mathbf{H}_{\mathrm{r}}$) ont été déterminés empiriquement après étude de la taille des objets et de la dispersion des données sur les différentes composantes de l'espace Lab. Les mêmes matrices échelles spatial et amplitude ont été utilisées dans ces expérimentations. Ceci permet de se focaliser sur l'influence du paramétrage des échelles temporelles h_{t^-} and h_{t^+}, afin d'obtenir

des résultats comparables à ceux obtenus par Gu *et al.* et Paris [50, 51].

La figure 9.2 montre les résultats des groupements obtenus avec trois combinaisons de h_{t^-} et h_{t^+}, permettant de réaliser un filtrage causal ou omniscient. Dans un premier temps, nous pouvons remarquer que quel que soit le réglage de (h_{t^-}, h_{t^+}) les objets sont bien segmentés. Cependant, la balle étant perdue suite à l'occultation par le rectangle lorsque seule l'image précédente est utilisée pour le filtrage ($h_{t^-} = -1$), une nouvelle classe est créée lors de sa réapparition à l'image 31. Au contraire, lorsque l'échelle temporelle autorise de prendre en considération jusqu'à trois images dans le passé ($h_{t^-} = -3$), la balle ne perd pas sa classe initiale après occultation. Ce résultat met en avant la capacité de la méthode STM-S^{++} à surmonter les occultations totales et temporaires d'objets par un réglage approprié des paramètre h_{t^-} et h_{t^+}. L'utilisation d'échelles temporelles plus grandes réduit les artéfacts de régions comme l'ombre de la balle et les contours de la raquette. Afin d'améliorer l'homogénéité des classes obtenus, un filtrage médian des images de classes ou tout autre algorithme de suppression de petites régions pourrait être utilisé.

9.3.2 Séquence football

L'approche STM-S^{++} a aussi été testée sur une séquence vidéo de football en accès libre sur internet [1]. Quatre images ont été extraites sur une seconde de capture vidéo afin que le maillot d'un joueur soit complètement masqué par un joueur adverse (Fig.9.3). Le pré-traitement appliqué est le même que pour la séquence *tennis de table*.

Les résultats obtenus par STM-S^{++} sont exposés Fig. 9.3. Bien que le résultat du groupement des images soit cohérent avec les données initiales, des régions comme le terrain de jeu ou les pantalons des joueurs sont sur-segmentés. De plus, le numéro cinq figurant sur le maillot du joueur change de classe lorsque uniquement l'image précédente est considérée dans le processus de filtrage. Il en est de même pour la classe associée au maillot du joueur rouge après son occultation. Lorsque les trois images précédentes sont prises en compte ($h_{t^-} = -3$), la classe associée au terrain n'est plus divisée en deux et les nombres des joueurs, leurs pantalons et la neige sont mieux identifiés. La cohérence temporelle des groupes est préservée relativement à l'échelle choisie et le maillot rouge ne change pas de groupe après occultation.

Néanmoins, nous pouvons remarquer que la classe du numéro cinq change à la dernière image. Cela vient d'un compromis "espace-temps". Il faut choisir entre une échelle spatiale dont la valeur sera proportionnelle au déplacement des objets pour pouvoir compenser un déplacement pendant une occultation et risquer de lier deux classes différentes mais similaires en amplitude et en temps, trop proches spatialement pour être différenciées. Dans ce cas, nous avons choisi une échelle spatiale importante pour compenser le déplacement des joueurs mais le chiffre cinq a été groupé avec le fond quand le joueur bleu s'est tourné, rendant ces deux classes similaires en amplitude trop proches spatialement pour être différenciées.

1. Séquence football : ftp://ftp.tnt.uni-hannover.de/pub/svc/testsequences/

9.4 Conclusion

Dans ce chapitre nous avons présenté une nouvelle formulation de mean-shift, STM-S^{++}, permettant d'unifier les approches de type causale et omnisciente. Nous avons montré le bon fonctionnement de cette approche sur deux séquences vidéos avec des occultations totales d'objets. Ainsi, la versatilité du formalisme proposé permet d'envisager son extension au filtrage de séries temporelles.

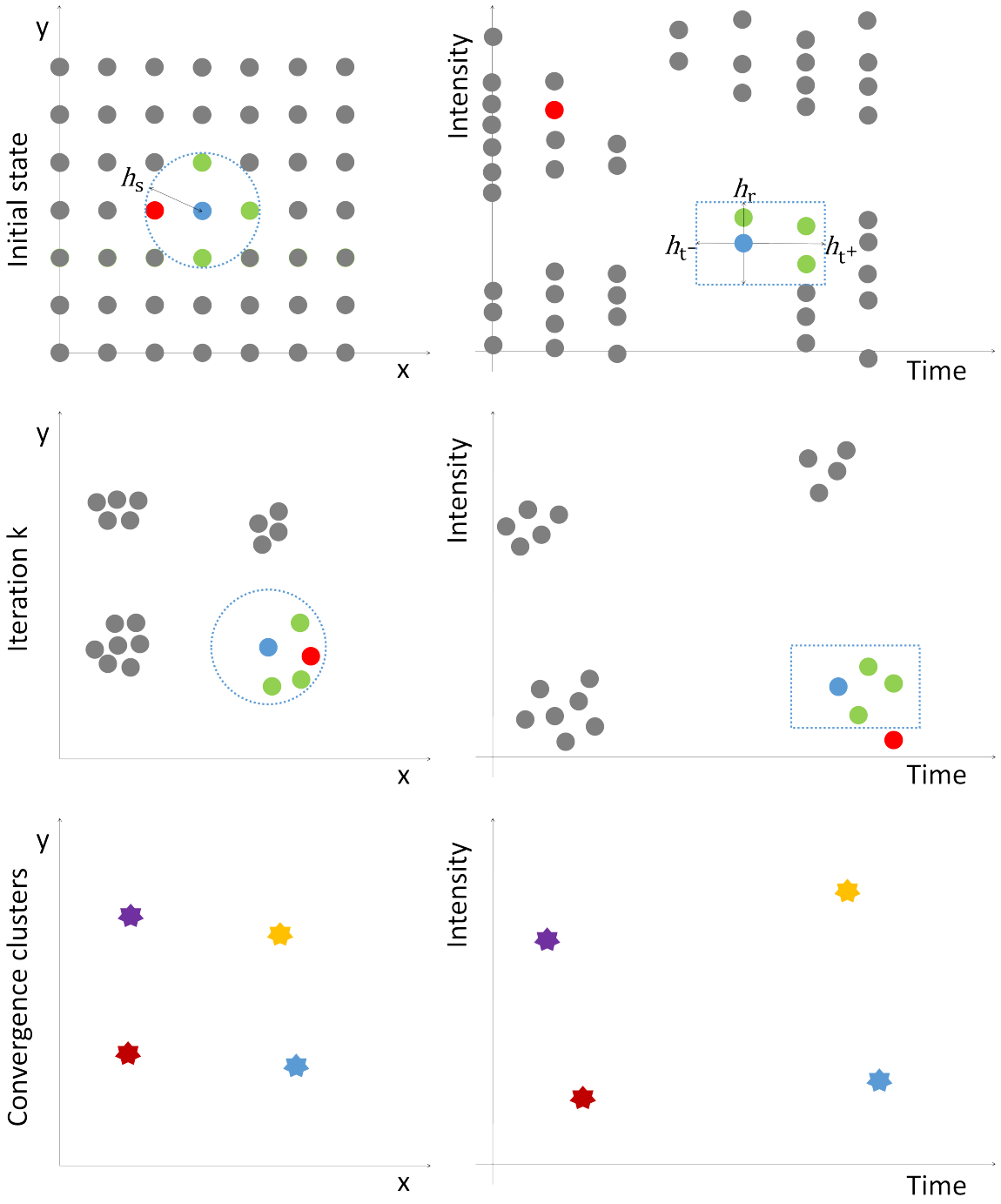


FIGURE 9.1 – Illustration de la convergence de STM-S^{++}. A gauche : support spatial. A droite : Supports amplitude et temporel. Les pointillés indiquent les voisinages sur les supports spatial, amplitude et temporel. L'observation bleue est celle utilisée comme référence à chaque itération. Les observations rouges sont les candidats pour mettre à jour la référence, mais qui ne sont pas à la fois compris dans le voisinage spatial, amplitude et temporel de la référence (strictement entre les pointillées). Les observations vertes sont les candidats qui satisfont ces trois conditions et servent à calculer le barycentre pour mettre à jour l'échantillon référence. Le nombre d'observations diminue au fil des itérations en utilisant l'algorithme de groupement conjointement au filtrage. On remarque que la composition de noyaux par multiplication génère un voisinage rectangulaire dans l'espace des caractéristiques.

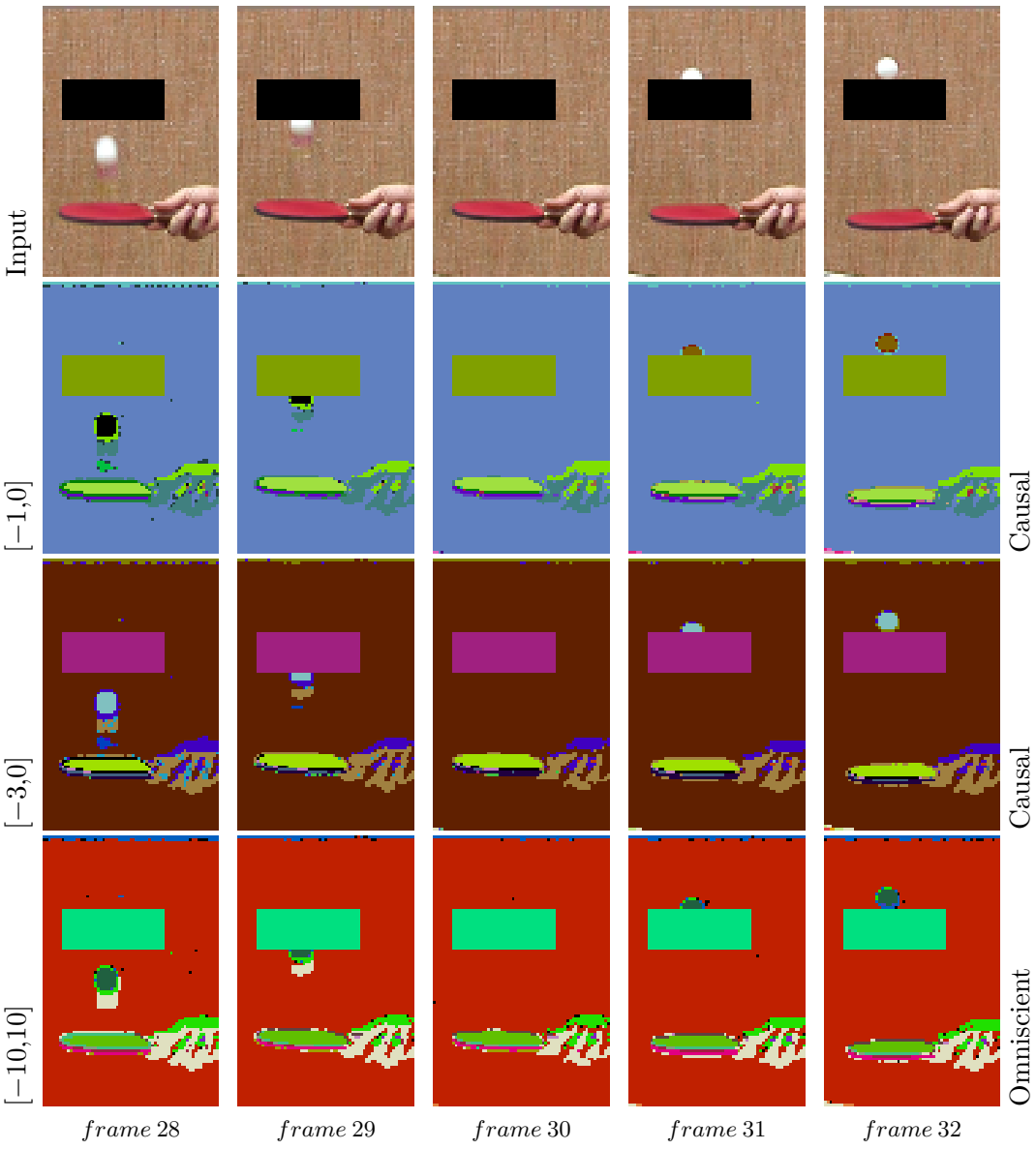


FIGURE 9.2 – Résultats obtenus sur la séquence tennis de table. La première ligne montre les images originales avec l'occultation. Les lignes suivantes montrent les résultats des groupements obtenus en utilisant les échelles temporelles spécifiées à chaque début de ligne. Les tests ont été effectués avec $\mathbf{H}_s = diag(40, 40)$ et $\mathbf{H}_r = diag(50, 30, 30)$. Aucun algortihme de suppression de petites régions n'a été appliqué.

FIGURE 9.3 – Résultats obtenus sur la séquence football. La première ligne montre les images originales. Les lignes suivantes montrent les résultats des groupements obtenus avec $\mathbf{H}_s = diag(20, 20)$, $\mathbf{H}_r = diag(10, 10, 10)$. Les échelles temporelles utilisées sont $h_{t^-} = -1, h_{t^+} = 0$ pour la deuxième ligne et $h_{t^-} = -3, h_{t^+} = 0$ pour la dernière ligne.

Chapitre 10

Conclusion

Les trois contributions méthodologiques majeures de ces travaux de thèse ont été présentées dans cette partie. Nous avons d'abord étendu le formalisme mean-shift au filtrage de séries temporelles et montré son intérêt par rapport à des méthodes de filtrage et de groupement existantes. Ces premier travaux ont donné lieu à la publication d'un article de journal dans la revue Pattern Recognition Letters et au colloque GRETSI 2015 [54, 60].

Nous avons ensuite proposé une méthode de groupement permettant d'identifier les séries temporelles similaires et déphasées en s'appuyant sur la mesure DTW. Ces travaux ont abouti à deux publications dans les conférences internationales ICASSP et ISBI [61, 62].

Pour terminer, nous avons étendu le formalisme M-S afin de prendre en compte explicitement les valeurs temporelles des échantillons. Ainsi, au lieu de traiter des échantillons indépendants il faudra traiter des ensembles d'échantillons comme ceux définis par (9.1) pour envisager le filtrage de séries temporelles ayant des fréquences d'échantillonnage différentes. Pour ce faire, les deux verrous méthodologiques à lever sont : (1) le choix d'une métrique de similarité des séries temporelles ; (2) le choix d'une stratégie de calcul de moyenne des séries temporelles. Ces travaux ont abouti a une publication dans la conférence internationale ICASSP [63].

L'originalité des méthodes proposées est qu'aucun a priori n'est fait sur le nombre de classes à trouver, ni sur leur structure ou sur un nombre minimum d'échantillons voisins. Ceux qui se ressemblent, relativement aux paramètres d'échelles choisis, convergent naturellement vers leurs représentants les plus probables.

Dans la partie suivante, nous appliquons ces méthodes dans le contexte de l'imagerie IRM de la sclérose en plaques afin d'identifier des ensembles de pixels dont les caractéristiques partagent des signatures temporelles similaires.

IV Applications des contributions méthodologiques aux données IRM de sclérose en plaques

Chapitre 11

Introduction

Dans cette partie, nous nous intéressons aux résultats applicatifs des algorithmes présentés en partie III dans le contexte de l'imagerie clinique de la sclérose en plaques (SEP). Nous étudions la capacité de l'approche STM-S à discriminer les différentes structures présentes dans les lésions de SEP (section 13.3.1).

Ensuite, nous utilisons l'algorithme de groupement hiérarchique basé sur la DTW pour étudier les similarités entre évolutions des différentes régions, que ce soit entre les lésions de SEP d'un même patient ou celles de plusieurs patients (section 13.3.2). Le but est d'étudier la possibilité de l'existence de sous-groupes de lésions partagés ou non entre les patients.

Pour terminer, nous étudions plus particulièrement avec STM-S les interactions des lésions SEP avec les veinules cérébrales au moment de leur apparition (section 14).

Chapitre 12

Imagerie par résonance magnétique et sclérose en plaques

12.1 Introduction

Dans les sections qui suivent, nous présentons les principes physiques de l'acquisition d'images par résonance magnétique (section 12.2). Ensuite, nous introduisons en section 12.3 ce qu'est la sclérose en plaques ainsi que le contexte de recherche environnant et les enjeux qui y sont liés.

12.2 L'imagerie par résonance magnétique

L'imagerie par résonance magnétique (IRM) est une technique d'imagerie non irradiante, permettant l'acquisition d'images anatomiques des tissus. C'est la technique de référence pour l'imagerie des tissus mous. Sa sensibilité à l'eau lui permet de visualiser des informations telles que la perfusion des organes (eau en mouvement), les inflammations (œudèmes) ainsi que la nature des tissus. Dans la plupart des cas, l'IRM est l'imagerie de l'atome d'hydrogène qui représente 63% des atomes du corps humain. Cette technique d'imagerie est fondée sur le principe de la résonance magnétique nucléaire (RMN) du noyau de l'atome d'hydrogène : le proton (spin nucléaire du proton). Un moment magnétique, appelé moment magnétique de spin, peut être associé au noyau d'hydrogène. Celui-ci peut être vu comme un petit aimant attaché au noyau d'hydrogène. Le principe de l'IRM est de détecter l'aimantation résultante de tous ces moments magnétiques. Cependant, les moments magnétiques ne sont pas orientés dans le même sens : leur distribution est isotrope.

Le champ résultant, appelé aimantation macroscopique est donc nul en l'absence de champ magnétique extérieur. Trois éléments constituant le système d'acquisition (Fig. 12.1) sont nécessaires pour obtenir une image en IRM.

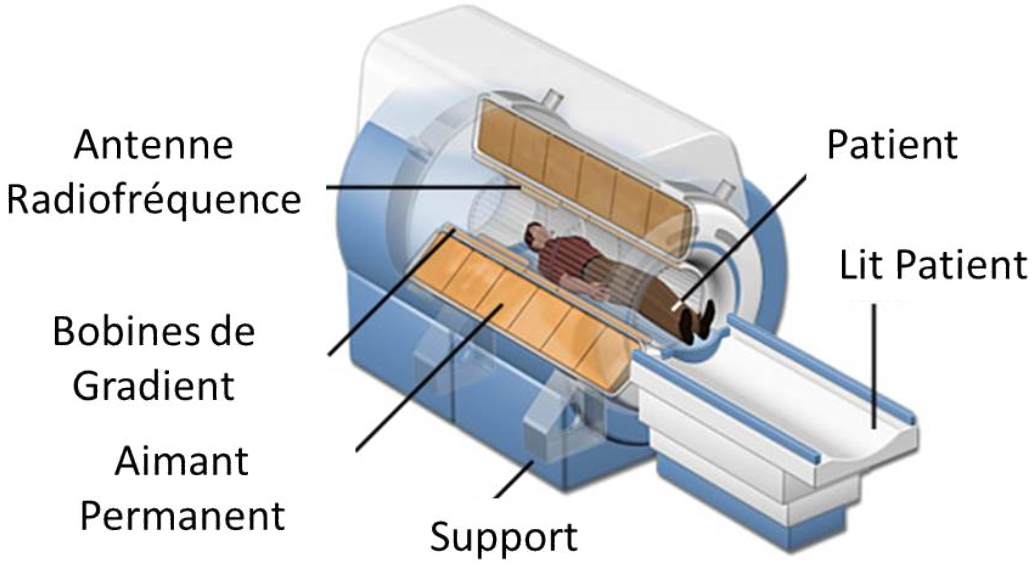


FIGURE 12.1 – Représentation schématique d'un appareil IRM, Coyne [64]. L'aimant permanent est supraconducteur, il baigne dans un bain d'hélium (-269°C) isolé de l'extérieur par un bain d'azote, des écrans thermiques et du vide.

Le premier est un aimant permanent appliquant un champ magnétique statique B_0. Une fois le patient placé dans l'appareil, cet aimant oriente les moments magnétiques portés par les molécules d'eau dans l'axe défini par le champ magnétique B_0 de l'aimant. Les moments magnétiques entrent ensuite en précession (Fig. 12.2), *i.e.* tournent comme des toupies autour de l'axe défini par B_0 à la fréquence de Larmor $\omega_0 = \gamma B_0$, avec γ le rapport gyromagnétique des protons.

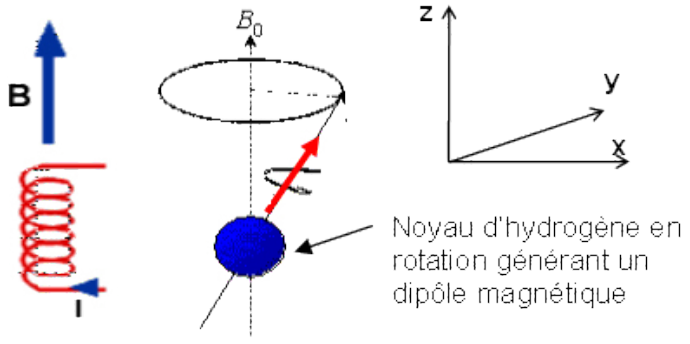


FIGURE 12.2 – En rouge : Aimant permanent permettant de générer le champ magnétique B_0. En bleu : noyau d'hydrogène/moment magnétique de spin en précession autour de B_0 à la fréquence de Larmor.
source : www.irmcardiaque.com

Un autre élément nécessaire est une antenne radiofréquence utilisée pour émettre et

réceptionner du signal (Fig. 12.3) : un champ radiofréquence B_1 est généré pour faire basculer l'aimantation macroscopique d'un certain angle. Cette impulsion radiofréquence fait basculer l'aimantation macroscopique d'un angle $\alpha = \gamma B_1 \tau$ dans le plan transverse à B_0, avec B_1 le champ radiofréquence et τ le temps d'application du champ/gradient. B_1 a la même fréquence que la fréquence de rotation des protons, ce qui donne l'énergie nécessaire à leur basculement : c'est le phénomène de résonance magnétique nucléaire. Une fois l'onde radiofréquence émise les protons reviennent à leur état initial, dans le sens de B_0. La composante dans le plan transverse diminue progressivement jusqu'à s'annuler tandis que la composante longitudinale repousse (dans le sens de B_0).

Le phénomène de relaxation se fait toujours en précession autour de B_0, ce qui permet à l'antenne de capter un signal RMN en sinusoïde amortie avec une pulsation égale à la fréquence de Larmor ω_0 et un amortissement égal au temps de relaxation sur l'axe transversal. Le signal RMN capté est aussi appelé Free induction Decay (FID) et il est caractérisé par son temps de relaxation transversal T_2 (temps que met le signal pour décroitre de 63% de sa valeur maximale dans le plan transverse) et son temps de repousse longitudinale T_1 ($T_1 > T_2$). Le principe de relaxation est illustré en Fig. 12.3.

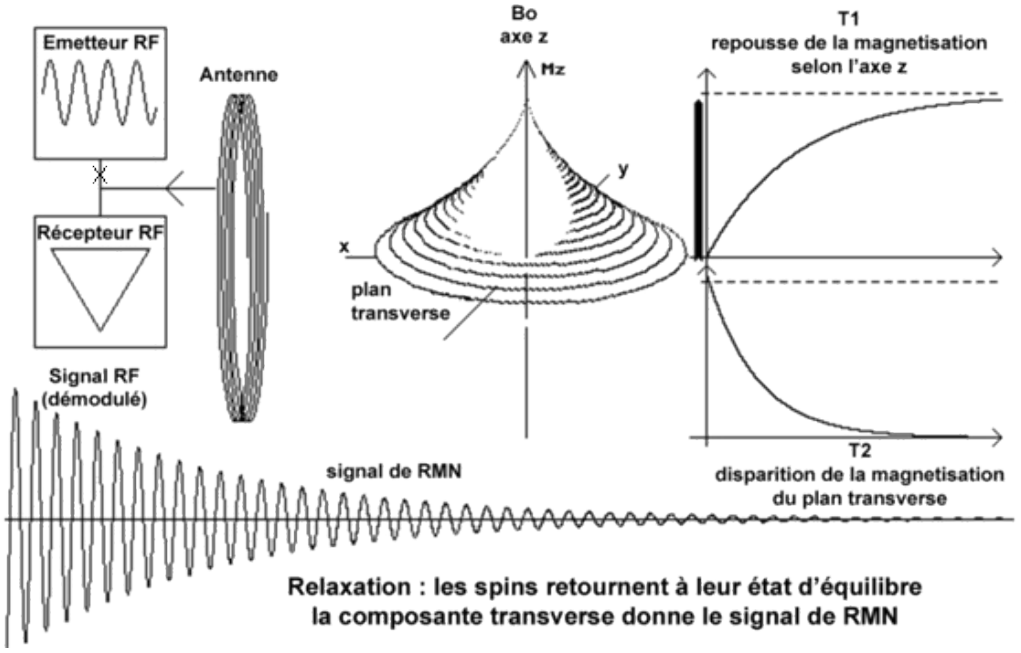


FIGURE 12.3 – Illustration du principe de relaxation du moment magnétique.
source : www.irmcardiaque.com

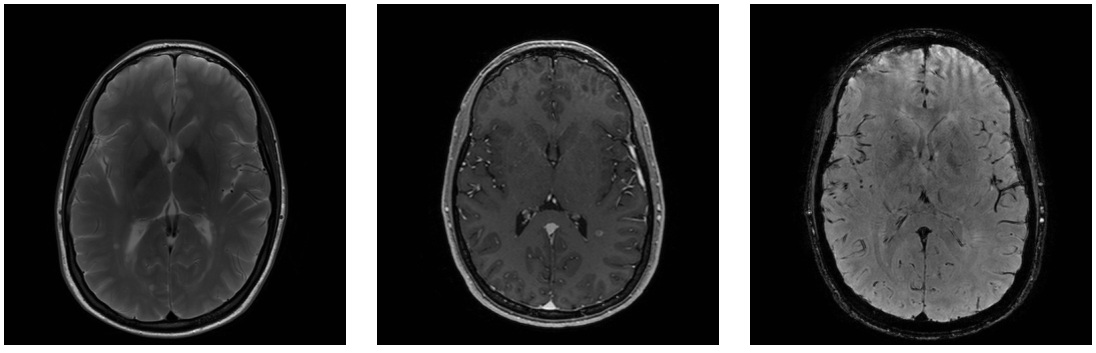
L'intensité du signal capté par l'antenne dépend du temps mis par le proton pour revenir dans l'axe de l'aimant (relaxation T1) et du temps mis pour se déphaser à nouveau (relaxation T2). Elle dépend aussi de l'instant auquel on l'acquiert après l'application de la radio fréquence : c'est le temps d'écho TE. Pour favoriser le contraste, il est plus intéressant d'utiliser des TE courts en T1 et longs en T2. Il faut savoir que chaque type de tissu en fonction de sa nature cellulaire et de sa micro-architecture comporte un T1 et un T2 qui lui sont propres. Ainsi, les images morphologiques obtenues pondérées en T1 et en T2

permettent d'analyser la nature normale ou pathologique d'un tissu précis.

Les derniers éléments constituant un système d'acquisition IRM sont les bobines de gradient de champ. Celles-ci servent à encoder le signal dans les trois directions de l'espace : la reconstruction d'une image est obtenue grâce à l'application de gradients de champ magnétique sur les axes x, y et z de l'espace (x,y : gradients de codage de phase et de codage de fréquence. z : gradient de sélection de coupe). Ceux-ci permettent d'encoder l'espace en fréquence ou en phase dans les trois direction de l'espace, car la fréquence de précession est directement reliée à la valeur du champ par le biais du rapport gyromagnétique. L'image obtenue est matricielle et présente un nombre fini de pixels (en 2D) ou de voxels (en 3D).

La qualité d'une image IRM dépend de nombreux paramètres : la résolution spatiale, le contraste, le bruit de la chaîne d'acquisition, le temps d'acquisition, l'intensité du champ statique, l'intensité des gradients, la qualité des antennes (homogénéité et sensibilité) et la présence ou non d'artéfacts. Les techniques d'imagerie en IRM permettent de créer des contrastes endogènes suivant la micro-architecture du tissu, en détectant à la fois l'interaction des molécules d'eau du tissu entre elles et avec le milieu environnant (champ magnétique statique, gradient de champ, graisse, interfaces tissulaires). Un contraste exogène peut aussi être créé par injection d'un agent de contraste qui va venir modifier localement le contraste endogène (en positif par renforcement du signal ou en négatif par atténuation du signal suivant les paramètres de la séquence). Généralement, ce produit de contraste est à base de gadolinium chélaté.

Nous utilisons ces différentes images, appelées séquences IRM, pour l'étude de la sclérose en plaques comme des séquences T2, T2 pondérées en phase (SWI) ou T1 avec injection de gadolinium pour renforcer le contraste de certaines lésions. Une illustration de trois types de contrastes obtenus avec trois séquences IRM différentes est visible Fig. 12.4. Dans la section suivante, nous présentons la sclérose en plaques.



(a) (b) (c)

FIGURE 12.4 – Différents contrastes obtenus en IRM. (a) Image pondérée en T2, l'eau apparaît hyper-intense (intensité réhaussée). (b) Image pondérée en T1 avec injection de gadolinium, les structures vasculaires et certaines lésions SEP apparaissent hyper-intenses. (c) Image de susceptibilité magnétique (SWI), les structures vasculaires apparaissent hypo-intenses (intensité atténuée).

12.3 La sclérose en plaques

La sclérose en plaques (SEP) est une maladie auto-immune, c'est à dire que le système immunitaire attaque ses propres cellules. Cette pathologie attaque plus particulièrement les cellules chargées de synthétiser la gaine de myéline qui entoure les axones dans le système nerveux central (cerveau et moelle épinière, cf. Compston [65]). Ce phénomène engendre des lésions dispersées dans le système nerveux central dont l'aspect est scléreux, *i.e.* épais et dur. Ces lésions appelées plaques ont donné leur nom à la maladie et sont le fruit d'une démyélinisation illustrée Fig. 12.5. Elles apparaissent par poussées inflammatoires entrainant le démyélinisation et sont souvent à l'origine d'une dégénérescence axonale. Un axone est l'unique lien par lequel un neurone communique avec sa cellule cible. La myéline quand à elle joue le rôle de gaine protectrice, c'est une membrane biologique présente autour des axones. Vecteur de l'information le long des neurones, la myéline est aussi impliquée dans la vitesse de propagation de l'influx nerveux.

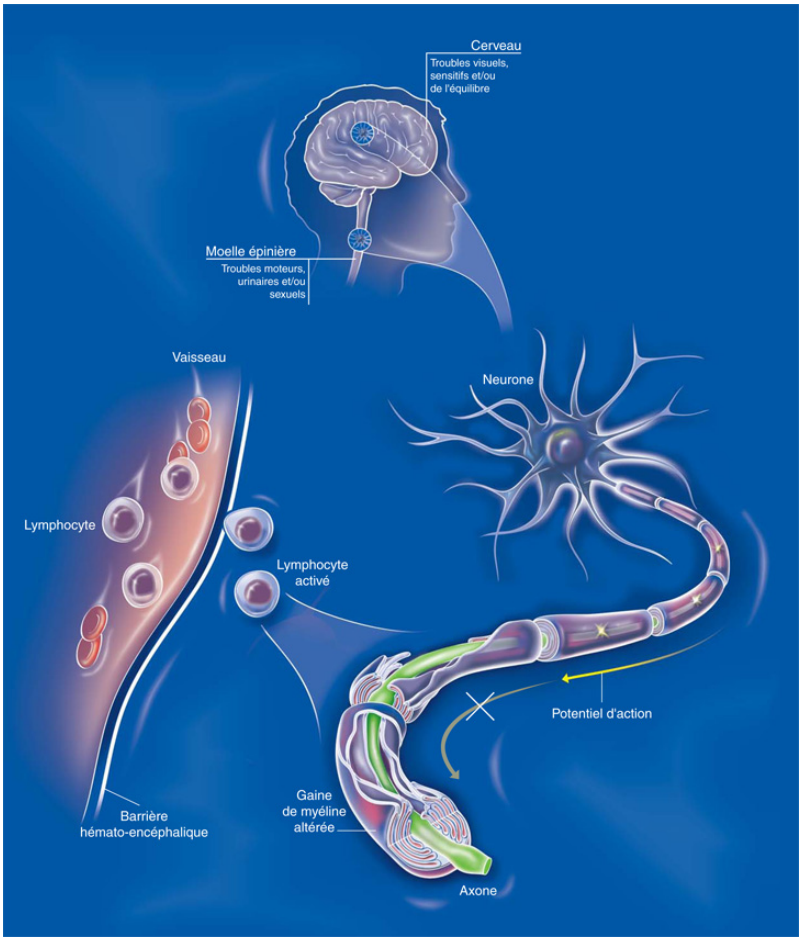


FIGURE 12.5 – Les lymphocytes attaquent et endommagent les gaines de myéline, ce qui perturbe ou empêche la circulation de l'influx nerveux dans les axones. *source : http ://www.inserm.fr/thematiques/neurosciences-sciences-cognitives-neurologie-psychiatrie/dossiers-d-information/sclerose-en-plaques-sep*

La sclérose en plaques est un problème de santé publique majeur. En France, elle touche environ 80000 personnes. Les causes peuvent être variables dans le temps et entre

les individus et aucun traitement efficace permettant de la soigner n'est connu à l'heure actuelle. Les facteurs génétiques et environnementaux déclenchant et guidant l'évolution de la pathologie, tout comme l'apparition de nouvelles lésions, sont encore très méconnus et représentent des domaines de recherche très actifs.

Afin de comprendre la génèse de cette pathologie, il est nécessaire d'étudier les évolutions et les interactions de ses facteurs endogènes (moléculaires, cellulaires) et exogènes qui aboutissent à la formation de nouvelles plaques et aux symptômes physiques affectant les patients. De fait, la recherche sur la génèse de la SEP nécessite des études multi-échelles allant des signatures temporelles des fluctuations de bio-marqueurs à la distribution spatio-temporelle des lésions dans le système nerveux central.

Au regard des lésions du système nerveux central, des analyses chimiques très sophistiquées et détaillées ont été menées sur des données statiques de biopsies ou d'autopsies de patients atteints de SEP [66]. Cependant, ces observations microscopiques ont été faites sur des données statiques. Par ailleurs, des travaux étudiant les comportements des lésions SEP à une échelle macroscopique, *e.g.* sur des séries d'images IRM, ont permis la modélisation de lésions SEP [65–68] via des analyses considérant les signatures temporelles des signaux dans les voxels [54, 69]. Pour terminer, en remontant d'un étage macroscopique, les évolutions spatio-temporelles des lésions ont été étudiées dans le contexte de la neuroanatomie fonctionelle pour comprendre le lien existant entre la formation de lésions et les symptômes perçus chez les patients [70], ainsi que les facteurs environnementaux affectant la saisonnalité de la prévalence de la SEP [71].

A l'heure actuelle, la caractérisation des lésions est majoritairement basée sur de l'information qualitative plutôt que de l'information quantitative extraite sur des données IRM : l'identification des sous-types de lésions est généralement faite par des experts évaluant leur apparence, guidés par les différentes séquences IRM disponibles (T2, T1 pre et post injection de gadolinium...) Ainsi le cerveau est cartographié de façon binaire, les choix étant faits entre "zone saine" et "lésion". Par conséquent, il existe un fort besoin pour développer des méthodes d'analyse permettant l'extraction automatique et la quantification des caractéristiques des lésions, en combinant de manière optimale les mesures hétérogènes disponibles. Par ailleurs, la combinaison de données hétérogènes avec des méthodes adaptées au groupement de données temporelles permettrait de distinguer les différents sous-types de lésions et d'identifier les sous-régions composant les lésions SEP. Ces informations pourraient décrire au sein d'une lésion des zones ayant différentes compositions histologiques, *e.g.* inflammation, œdème, démyélinisation/dégénérescence axonale, remyélinisation, ce qui pourrait apporter de nouvelles perspectives sur la pathophysiologie de la SEP.

12.4 Conclusion

Dans ce chapitre, après avoir brièvement rappelé les principes d'acquisition d'images par résonance magnétique, nous avons présenté la sclérose en plaques, ses conséquences sur

le système nerveux et certaines questions encore ouvertes sur le sujet. Il reste cependant beaucoup de chemin à parcourir, tant sur la compréhension de cette pathologie que sur la découverte de traitements qui permettraient de la soigner. Dans le chapitre suivant, nous essayons d'amener des éléments de réponse en appliquant nos contributions méthodologiques à des séries d'images acquises en IRM sur des patients atteints de sclérose en plaques.

Chapitre 13

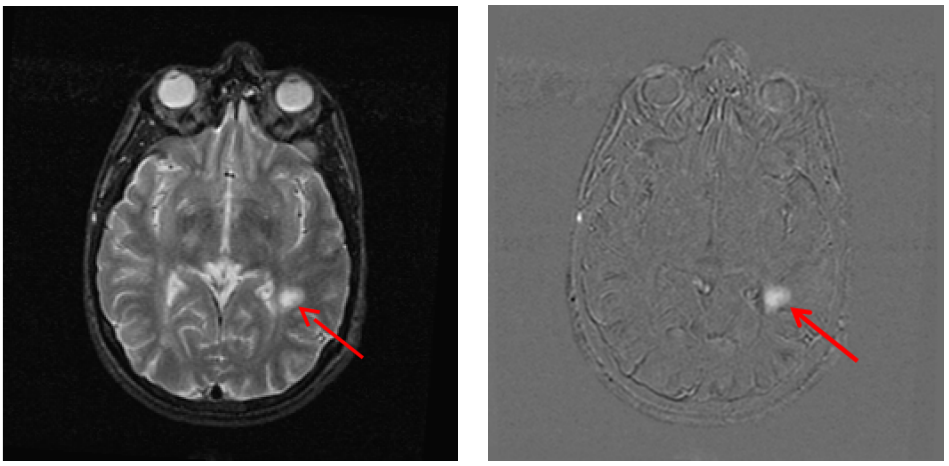
Etude de l'évolution des lésions de SEP en IRM

13.1 Introduction

Dans cette section, les travaux portent sur un jeu de données (Boston40) composé de 13 individus sélectionnés parmi une étude d'histoire naturelle [72] menée sur des patients affectés par la SEP de forme cyclique/récurrente-rémittente (Fig. 13.1). Nous présentons les données ainsi que leur chaîne de traitements en section 13.2. Les résultats obtenus avec STM-S sont présentés section 13.3.1 tandis que ceux obtenus ensuite avec notre algorithme de groupement utilisant la DTW sont présentés section 13.3.2. Le but de cette étude est de mettre en évidence l'existence de (dis)similarités entre les signatures temporelles des intensités des lésions de SEP en intra et en inter patients.

13.2 Matériel et méthode

Les patients contenus dans Boston40 ont été suivis via, en moyenne, 24 examens IRM répartis sur une année. Les huit premiers examens IRM ont été acquis toutes les semaines, les huits suivants toutes les deux semaines tandis que les derniers examens sont mensuels. La séquence IRM qui est utilisée dans cette étude est un protocole dual-echo spin-echo avec des coupes contigües de trois millimètres d'épaisseur permettant d'imager la totalité du cerveau (Fig. 13.1). Trente régions d'intérêt (ROI) 3D+t ont été sélectionnées, chacune contenant une nouvelle lésion SEP. Avant d'appliquer STM-S sur les ROI retenues, nous avons utilisé une chaîne de prétraitements inspirée par [68]. Dans un premier temps, nous



(a) (b)

FIGURE 13.1 – (a) Coupe d'une image IRM pondérée en T2 d'un patient atteint de SEP. (b) Image de différence entre l'image en (a) et la première acquisition du suivi, où aucune lésion n'est encore visible. La flèche rouge pointe sur une lésion SEP. Il s'agit de la même coupe, la lésion SEP est plus visible sur l'image de différence (b).

avons extrait la boîte crânienne dans chacune des acquisitions avec FSL BET [73]. Sans cette étape, le crâne est très bien recalé entre les séquences d'un même patient mais le cerveau n'est généralement pas bien aligné à cause de possibles déplacements dans la cavité crânienne. Ensuite, les distributions d'intensités des cerveaux ont été normalisées (centrées et réduites) de manière indépendante à chaque point temporel. Les séquences IRM de chaque patient ont été recalées sur la première acquisition de leur suivi. Nous avons utilisé un recalage rigide à 6 degrés de liberté avec le logiciel FSL FLIRT [74, 75] (Fonction de coût : moindres carrés. interpolation : sinus cardinal). La précision obtenue après recalage est au pixel près. Il s'agit d'un point très important dans la mesure où STM-S traite les évolutions des caractéristiques des voxels. Si le recalage n'est pas bon alors les séries temporelles traitées ne correspondraient pas aux évolutions réelles, ce qui fausserait les résultats. Pour terminer, la première séquence IRM de chaque suivi a été soustraite aux séquences suivantes. Le fait de considérer la première acquisition comme une référence, en travaillant avec les images de soustraction, permet de focaliser l'étude sur les nouvelles lésions SEP ou celles dont l'intensité évolue par rapport au temps zéro et d'annuler toutes les valeurs constantes au cours du temps.

Après avoir indépendamment prétraité les séquences IRM pondérées en T2 de chacun des suivis, trente ROI ont pu être extraites sur un total de treize patients. Chacune d'elles contient une nouvelle lésion apparaissant après la première acquisition et avant la seizième, ce qui permet de capturer les apparitions des lésions avec un échantillonnage inférieur ou égal à deux semaines. La répartition du nombre de ROI extraites par patient est illustrée Fig. 13.2.

L'approche STM-S présentée section 7.2, puis la méthode de groupement basée sur la DTW que nous avons détaillée section 8.2 (Fig. 13.3) sont utilisées pour le traitement des

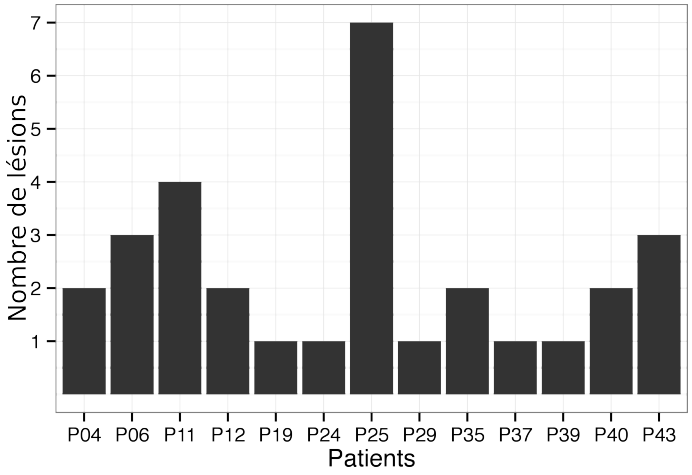


FIGURE 13.2 – Nombre de ROI par patient pour les treize patients considérés.

suivis longitudinaux de chaque patient. Cet enchainement permet dans un premier temps de trouver les groupes de pixels qui évoluent similairement dans chaque ROI. Ensuite, les évolutions d'intensités similaires de ces groupes peuvent être identifiées entre toutes les ROI, même si les signatures temporelles sont déphasées.

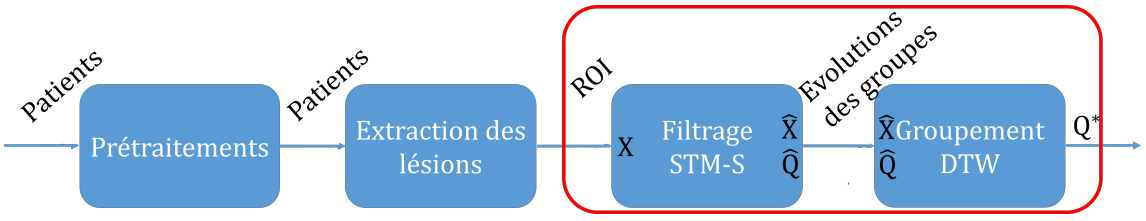


FIGURE 13.3 – Chaîne de traitements des données IRM. $\mathbf{X}, \hat{\mathbf{X}}, \hat{\mathbf{Q}}$ et $\mathbf{Q}^*$ ont été définis en section 8.2. L'encadrement rouge identifie nos contributions méthodologiques.

13.3 Résultats

13.3.1 Filtrage STM-S

Dans un premier temps, nous voulons déterminer la capacité de l'approche STM-S à discriminer différents groupes de pixels dans les ROI en se basant sur leurs évolutions de niveaux de gris. Nous voulons aussi montrer que STM-S permet non seulement de conserver les comportements intrinsèques des lésions, mais aussi de discriminer leurs différentes structures internes.

Le choix des paramètres a été déterminé de manière empirique en nous appuyant sur les résultats présentés section 7.3.3 : l'échelle spatiale h_s est égale à 1000 pixels afin de considérer tous les pixels dans les ROI comme des voisins potentiels tandis que l'échelle amplitude h_r est égale à 0.6. Il ne faut pas oublier que les intensités des images sont normalisées, elles sont donc majoritairement comprises entre les valeurs -1 et 1. Ces paramètres

sont identiques pour le traitement de toutes les ROI, hormis deux dont l'intensité maximale des lésions au cours du suivi n'est pas très élevée. Pour celles-ci, l'échelle h_r est fixée à 0.3 afin que les lésions puissent être distinguées du fond.

Les figures 13.4 et 13.5 montrent des exemples de résultats obtenus par notre approche sur des lésions SEP. Le filtrage STM-S des ROI permet de détecter les lésions et met aussi en évidence la forme concentrique de leurs évolutions, déjà observée par Guttmann *et al.* [72] et Meier *et al.* [67]. Nous arrivons aussi à identifier les structures internes des lésions : une partie centrale appelée le cœur et une partie périphérique appelée l'œdème. En regardant les signatures temporelles des intensités associées à ces parties, nous pouvons remarquer que les cœurs de lésions sont similaires en forme aux œdèmes mais ont un pic d'intensité plus important, correspondant à une inflammation plus conséquente. De plus, les évolutions temporelles représentatives des groupes de pixels identifiés par notre approche sont similaires, pour ceux que nous associons à des lésions, au modèle d'évolution d'intensité des lésions SEP proposé par Meier *et al.* [67]. Il est important de rappeler que les mêmes paramètres ont été utilisés pour traiter ces ROI, par conséquent notre méthode permet d'obtenir des résultats cohérents avec un réglage simple et intuitif de ses paramètres. Bien que STM-S ne permette pas de modéliser des régions dont la forme varie au cours du temps, il est intéressant de constater que ces premiers résultats restent en accord avec les précédentes études et permettent d'identifier différents compartiments tels que le cœur et l'œdème des lésions.

Nous venons de montrer sur notre jeu de données que l'approche STM-S fonctionne bien quand elle est utilisée indépendamment sur chaque ROI. Nous arrivons à filtrer et bien identifier des groupes de pixels dont les niveaux de gris évoluent de manière synchrone au cours du temps. Cette première étape franchie, nous souhaitons étudier si pour un même patient et entre patients différents les groupes de pixels identifiés comme des lésions évoluent de la même manière au cours du temps. Il est très peu probable que les lésions apparaissent et disparaissent en même temps chez un patient, cela l'est encore moins lorsque nous étudions plusieurs patients. Pour arriver à identifier les similarités qui pourraient exister entre les différentes lésions il est nécessaire d'utiliser une mesure qui ne soit pas sensible aux décalages temporels, ce que nous proposons dans la section suivante en utilisant l'algorithme présenté en section 13.3.2.

13.3.2 STM-S + groupement DTW

Après filtrage et groupement STM-S des 30 ROI extraites (Fig. 13.2), nous avons sélectionné les représentants, *i.e.* les évolutions moyennes, de chaque groupe de pixels appartenant à une lésion. Nous voulons associer les groupes dont les évolutions moyennes d'intensité sont similaires en restant insensibles à de potentielles distorsions temporelles induites par la nature asynchrone de la SEP. Nous disposons désormais de 75 évolutions, représentantes des groupes identifiés à l'étape précédente, que nous allons utiliser dans la procédure de groupement basée sur la DTW présentée section 8.2. Par conséquent, nous

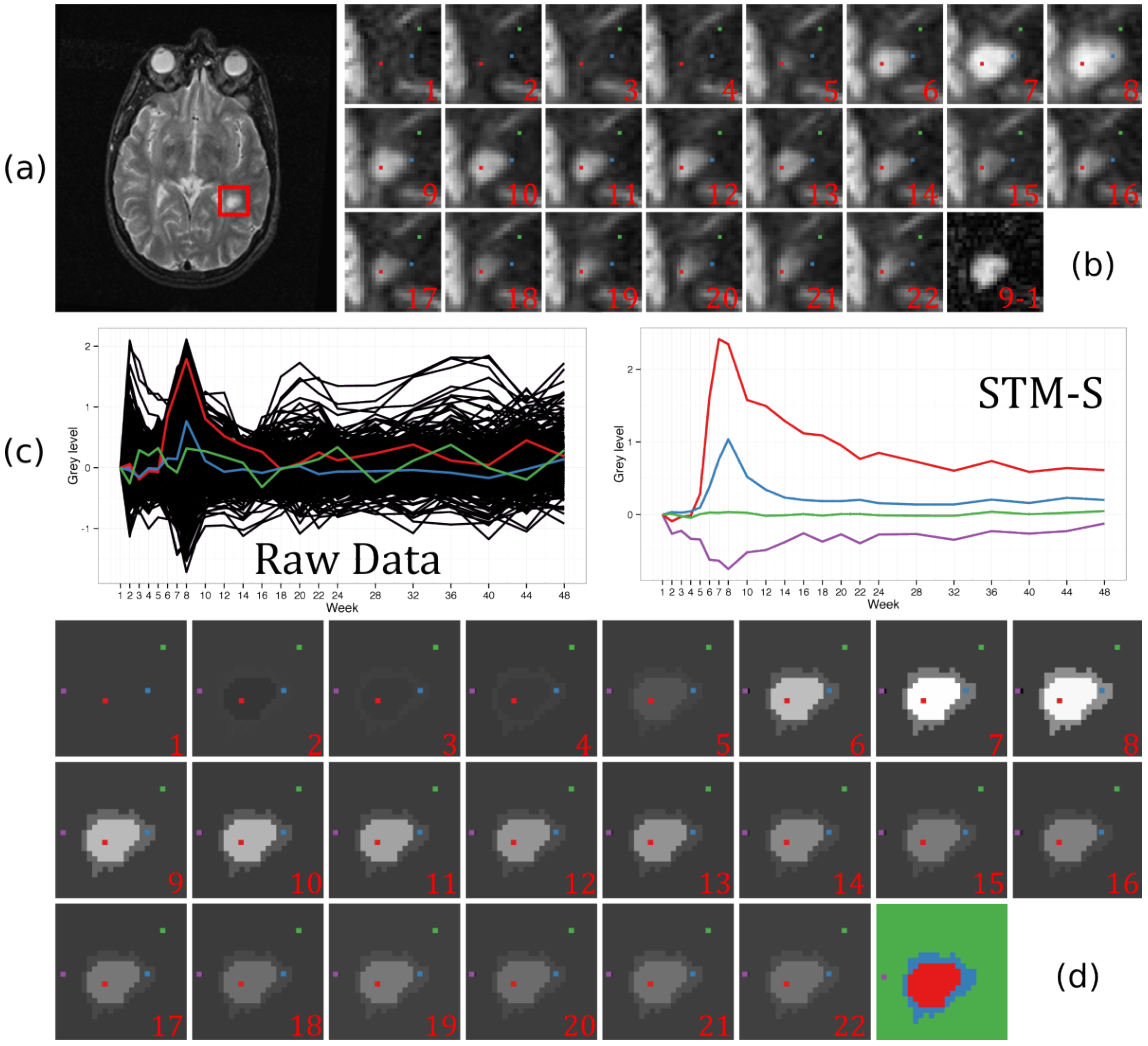
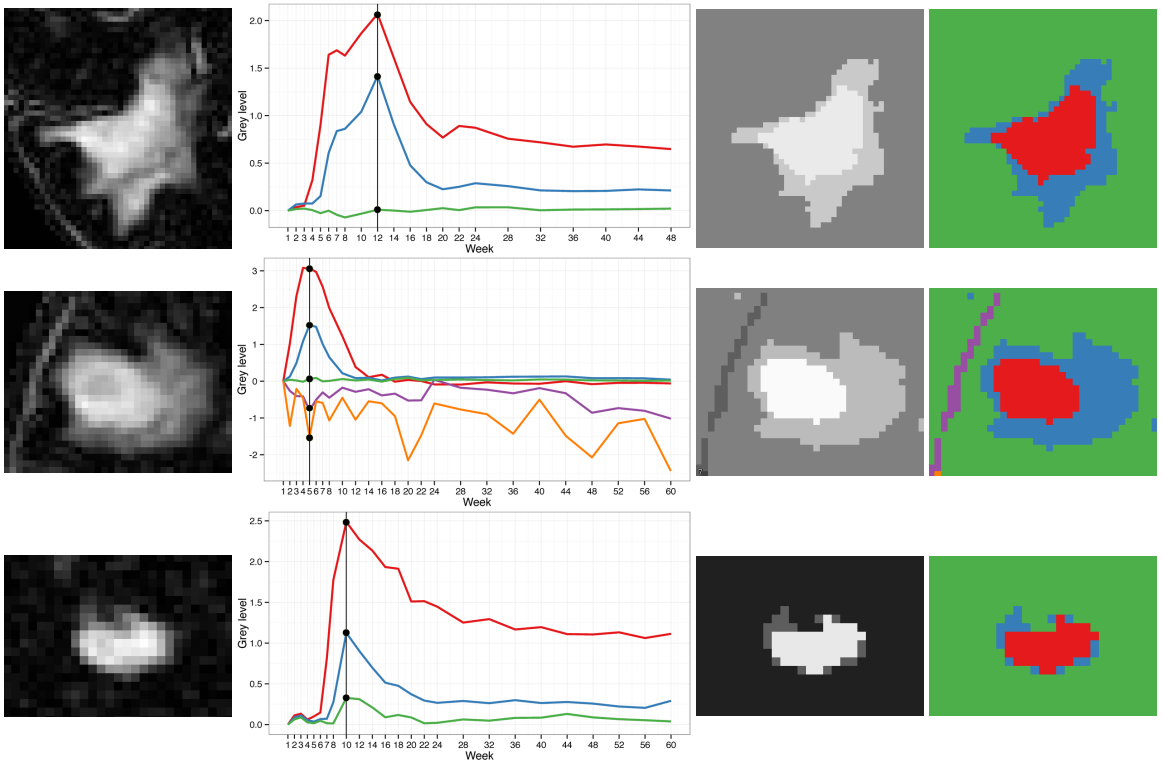


FIGURE 13.4 – (a) Coupe d'une acquisition IRM pondérée en T2 et identification de la ROI. (b) Evolution de la lésion présente dans la ROI au cours de 22 instants, la case 9-1 permet de voir l'image de soustraction du 9è point temporel. (c) A gauche : évolutions temporelles des intensités de tous les voxels de la ROI, les trois courbes en couleur correspondent aux évolutions des voxels d'intérêt identifiés en (b). A droite : évolutions temporelles des intensités de tous les voxels après groupement STM-S. (d) Résultat du filtrage STM-S de la ROI. En bas à droite : Groupes obtenus après filtrage STM-S de la ROI, les couleurs correspondent aux évolutions après groupement STM-S en (c).

voulons trouver les similarités entre les évolutions des lésions en intra et en inter-patients.

La table 13.1 montre la capacité de l'algorithme à trouver les similarités existantes entre les lésions appartenant à un même ou a plusieurs patients. Après filtrage et groupement des 30 ROI, STM-S a identifié 75 groupes de lésions différentes. Une fois traités par notre algorithme basé sur la DTW, 22 contiennent des évolutions n'étant pas partagées par d'autre lésion, 3 contiennent des évolutions initialement associées à différentes lésions d'un même patient et 13 contiennent des évolutions initialement associées à différentes lésions dans différents patients. Cela montre qu'il existe bien des similarités entre les évolutions



(a) (b) (c) (d)

FIGURE 13.5 – (a) Coupes de trois ROI 3D différentes englobant une lésion SEP. (b) Evolution des niveaux de gris après filtrage STM-S, avec $h_s = 1000$ and $h_r = 0.6$. (c) Filtrage STM-S de (a). (d) Groupes obtenus après filtrage STM-S de (a). Les couleurs des pixels en (d) sont associées aux évolutions de même couleur en (b).

des lésions de SEP en intra et en inter patients. On peut aussi voir Fig. 13.6 l'avantage du groupement basé sur la DTW par rapport à STM-S. Une même classe comporte des évolutions dont les formes sont similaires mais apparaissant à différents instants du temps. Par exemple, les évolutions en couleurs Fig. 13.6(a) ont la même forme et appartiennent au même groupe malgré leurs déphasages temporels. Pour terminer, les images de certains des groupes obtenus en Fig. 13.6(b) montrent les similarités structurelles que notre méthode permet de retrouver entre différentes lésions. Néanmoins, on peut voir Fig. 13.7 la limitation de l'algorithme de groupement proposé. En comparant les évolutions du groupe bleu et l'évolution du groupe marron de la lésion 3 du patient 25 (P25L3), on aurait pu espérer qu'elles soient toutes associées à la même classe, leurs formes étant toutes similaires. On peut suspecter l'aspect trop restrictif de notre contrainte de lien pour l'algorithme de groupement hiérarchique. La norme infini sur les écarts entre les mesure associées ne tolère pas le moindre dépassement du paramètre d'échelle. De plus, plus un groupe comprendra des séries temporelles et plus il sera difficile d'en associer de nouvelles car il faut qu'elles soient proches de tous les membres du groupe par rapport à (8.5).

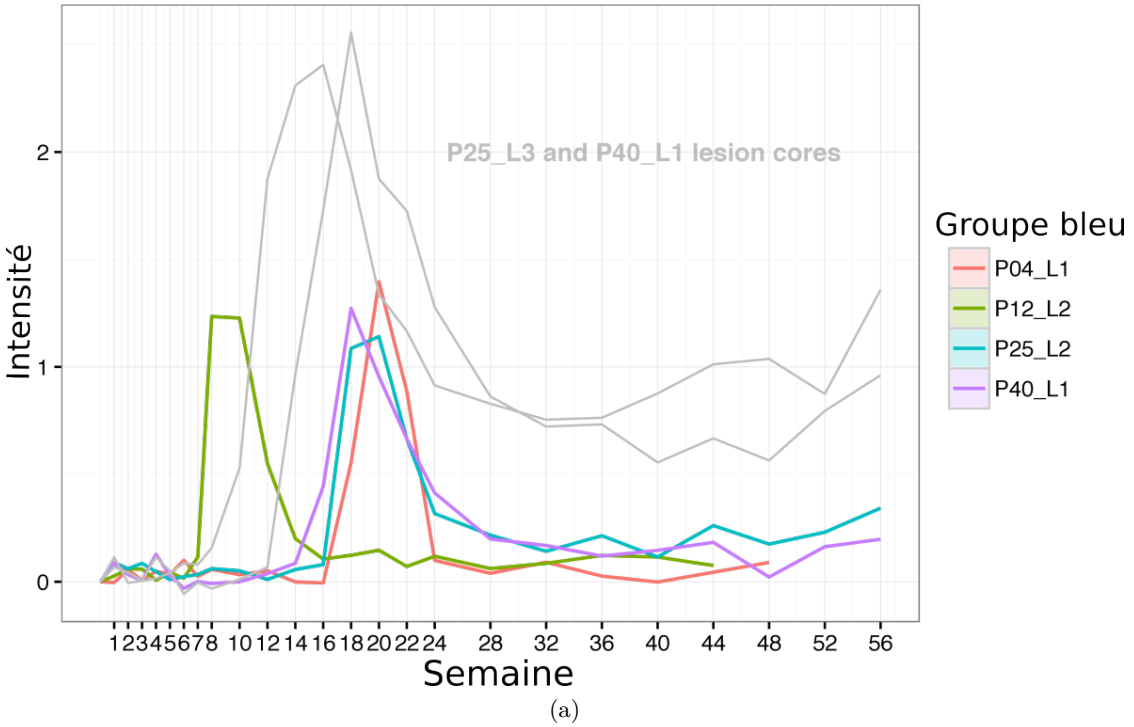
TABLE 13.1 – Impact du groupement DTW sur le nombre de groupes après filtrage STM-S (baisse de 75 à 38 groupes). Sur chaque ligne sont présentés le nombre de groupes seulement associés à une lésion, le nombre de groupes associés à plusieurs lésions dans un patient et le nombre de groupes associés à plusieurs lésions et partagés entre plusieurs patients.

	Nombre de groupes	
Groupes associées à	**Sans groupement DTW**	**Groupement DTW**
Une lésion / Un patient	75	22
Plusieurs lésions / Un patient	0	3
Plusieurs lésions / Plusieurs patients	0	13

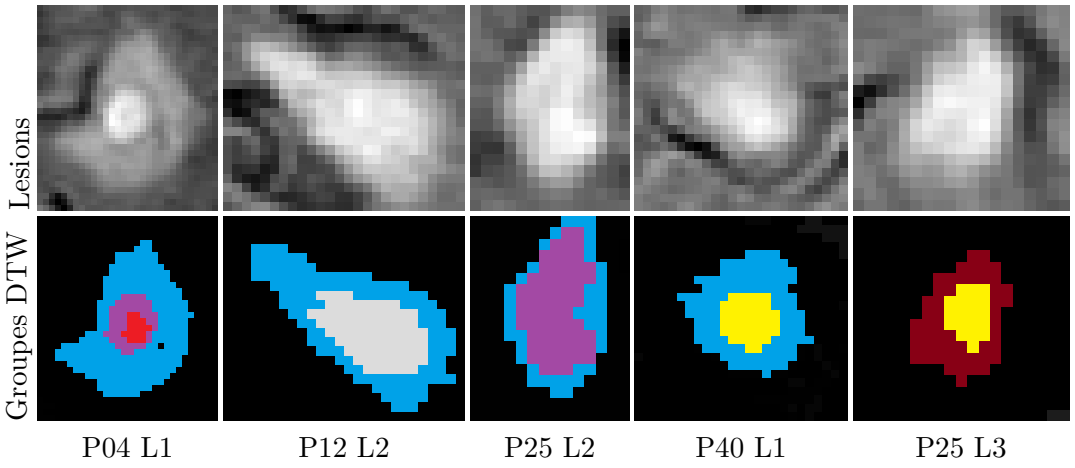
13.4 Conclusion

Les premiers tests effectués sur des données de sclérose en plaques sont encourageants et en accord avec les observations de la littérature [67]. Ce travail est un premier pas vers une étude plus approfondie de la classification des lésions de SEP en fonction de leurs comportements. Des comportements similaires ont été observés entre des lésions appartenant à différents patients. Néanmoins, l'algorithme de groupement basé sur la DTW est efficace en tant que preuve de concept mais risque de s'avérer plus limité pour des études à grande échelle de part sa contrainte de fusion des séries temporelles restrictive. En d'autres termes, plus la taille des groupes augmente et plus il sera difficile qu'un nouvel échantillon soit similaire à tous les membres d'un groupe. Il sera donc nécessaire d'utiliser un algorithme de groupement plus avancé afin de ne pas sur-estimer le nombre de groupes existants.

Après avoir abordé certains aspects macroscopiques de la SEP via des groupes de lésions et de patients, nous allons descendre d'une échelle dans le chapitre suivant pour étudier les interactions entre les lésions SEP et les veinules cérébrales. Nous nous intéressons à ce qui pourrait être un mécanisme à l'origine de l'apparition des lésions dans le cerveau, le phénomène de pincement des veinules cérébrales par les lésions SEP.



(a)



(b)

FIGURE 13.6 – Les nombres après P désignent les numéros des patients et les nombres après L désignent les numéros des lésions. (a) Exemple de deux classe obtenues après groupement DTW des ROI. Les évolutions en couleur appartiennent au groupe bleu rassemblant les œdèmes de quatre patients différents visibles en (b). Les courbes grises appartiennent aux groupes jaunes visibles en (b) correspondant à deux cœurs de lésions. (b) Coupes des ROI 3D avant et après groupement DTW. Les couleurs qui correspondent entre les lésions permettent d'identifier les classes identiques après groupement DTW.

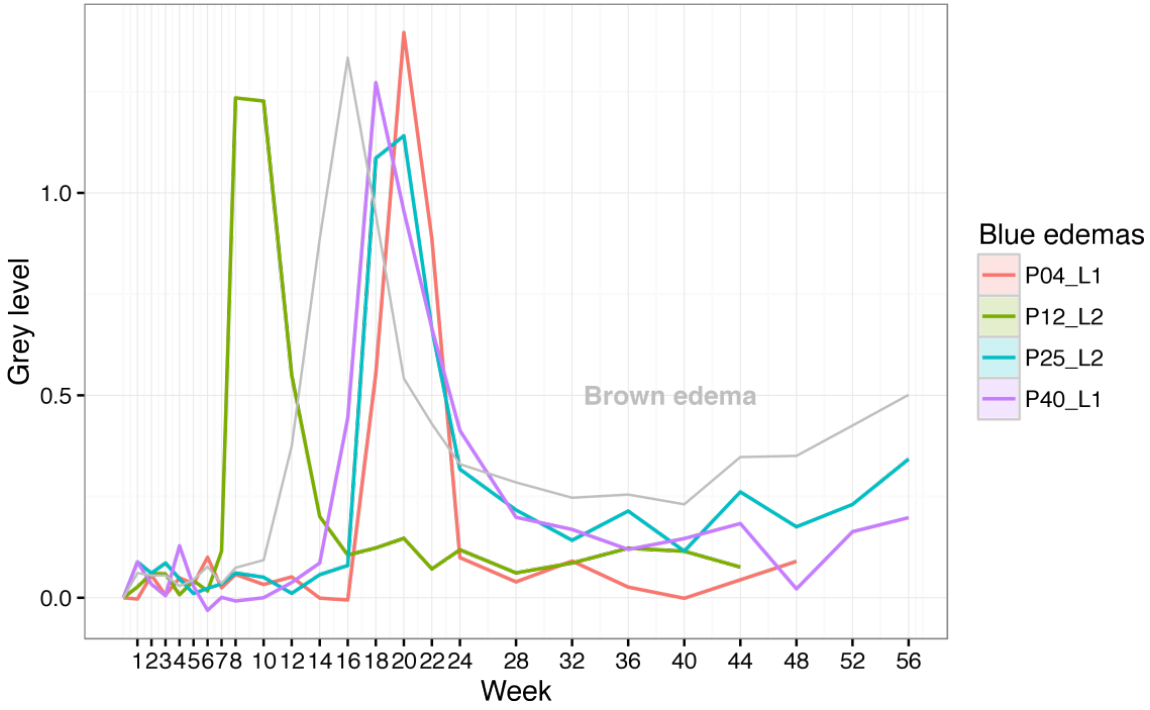


FIGURE 13.7 – Evolutions des groupes "œdème bleu" (en couleurs) et "œdème marron" (en gris). Ces classes n'ont pas été associées par le groupement basé sur la DTW.

Chapitre 14

Etude de l'interaction des lésions de SEP avec les veinules cérébrales

14.1 Introduction

De manière générale, les lésions SEP localisées dans la matière blanche du cerveau se développent autour d'une veine centrale [76]. Ce phénomène est connu depuis longtemps grâce à l'étude de données histologiques [77, 78]. Un développement centrifuge a tout d'abord été suspecté mais aucune donnée dynamique n'était alors disponible. Avec l'arrivée de l'IRM, des suivis longitudinaux permettant de vérifier cette dernière hypothèse [72] ont pu être acquis. Plus récemment, la modification simultanée du diamètre des veines cérébrales avec l'apparition des lésions SEP a été observée [79]. Ces résultats supposeraient que la petite taille apparente des veines intra lésionnelles puisse être causée par leur compression par un manchon inflammatoire à l'intérieur des lésions ou par un durcissement des parois vasculaires dans les lésions chroniques. Néanmoins, à notre connaissance aucune étude n'a porté sur les interactions existantes entre les veines et les lésions SEP lors de leur apparition, grâce à un jeu de données dynamiques. Nous proposons ici d'utiliser la séquence IRM de susceptibilité magnétique (SWI), permettant de mettre en valeur les veinules cérébrales en exploitant l'oxygénation sanguine [80]. Nous nous intéressons plus particulièrement à l'évolution de l'intensité des pixels localisés à l'interface des veines et des lésions SEP. Le but est d'identifier un possible pincement, appelé sténose, des veinules cérébrales lors de l'apparition des lésions SEP. Ce comportement spécifique est supposé en phase avec la prise de contraste des lésions en IRM pondérée T1 avec injection de

gadolinium (T1-gado, cf. section 12.2), témoignant de la naissance des lésions.

Dans le chapitre suivant, nous décrivons le jeu de données IRM sur lequel nous avons travaillé (section 14.2) avant d'exposer les résultats obtenus (section 14.3) et de conclure (section 14.4).

14.2 Matériel et méthode

Le jeu de données utilisé a été acquis de manière hebdomadaire pendant huit semaines sur cinq patients non traités atteints par la SEP de forme cyclique [81, 82]. La chaîne de de prétraitements utilisée est la même que dans le chapitre précédent (voir section 13.2), et a été appliquée indépendamment aux séquences T1-gado et SWI.

Les nouvelles lésions ont été identifiées par un expert, analysant conjointement les séquences T1-gado et SWI. C'est cette dernière modalité IRM qui est utilisée pour identifier les possibles changements d'apparence de la veine lors de l'apparition des lésions, le T1-gado ne servant qu'à l'identification des nouvelles lésions. Deux critères de sélection ont été utilisés pour extraire les lésions : (1) La prise de contraste des lésions doit apparaître après la première acquisition et disparaître avant la dernière, afin d'assurer un suivi complet de leur phase active. (2) Les veines sur lesquelles sont centrées les lésions doivent avoir un diamètre minimal de deux voxels dans le plan xy. Leurs dimensions sont de 0.344mm dans le plan xy et l'épaisseur de coupe est de 0.7mm.

Les ROI 3D extraites dans les IRM SWI ont été traitées avec l'approche STM-S dans le but de grouper les pixels ayant des évolutions d'intensité similaires et synchronisées. Par conséquent, aucun a priori sur le comportement recherché à l'interface veine/lésion n'a été fait pour initialiser et/ou guider le processus de groupement. Au final, sur 101 lésions remplissant le critère de prise de contraste, seulement trois appartenant à deux des cinq patients sont centrées sur des veinules suffisamment larges pour être observées. Nous présentons les résultats obtenus sur ces trois cas dans la section suivante.

14.3 Résultats

Nous pouvons observer Fig. 14.1 que les lésions sont bien extraites après le groupement avec STM-S. Pour chacune des lésions nous pouvons aussi remarquer qu'elles comportent une ou deux petites classes à l'intérieur. D'après la définition de la contrainte de proximité temporelle dans le formalisme de STM-S (7.8), nous pouvons assurer que ces groupes n'ont pas des comportements similaires aux autres groupes présents dans leurs ROI. Plus précisément, ces comportements particuliers sont des hyper-intensités transitoires synchrones à la prise de contraste en T1-gado, localisées à l'intérieur des veines sur lesquelles les lésions sont centrées. Les groupements obtenus permettent d'identifier plusieurs groupes à l'intérieur des lésions, dont un localisé à chaque fois en leurs cœurs entre les veines et la matière blanche (Fig. 14.1).

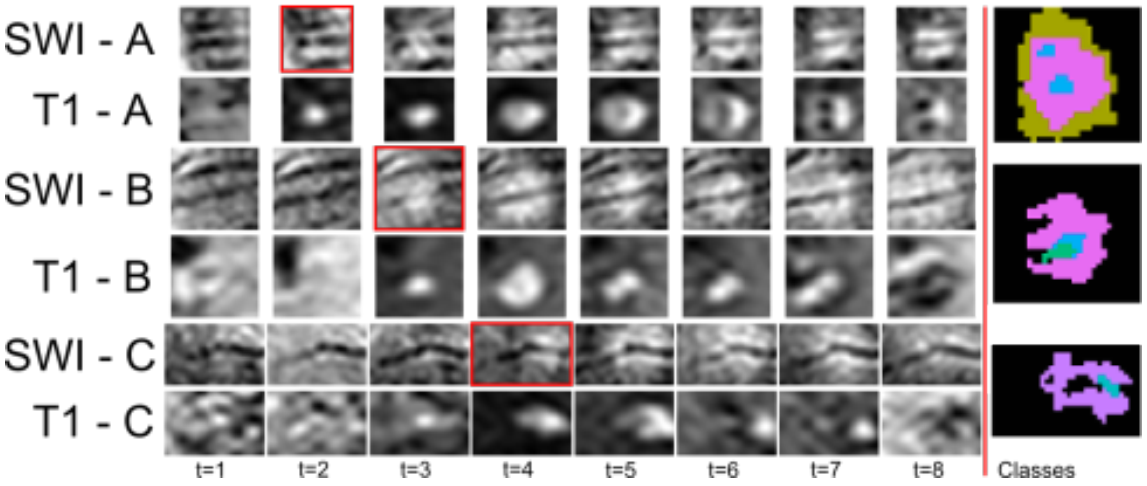


FIGURE 14.1 – 8 points temporels acquis pour les ROI (A, B, C) acquises en T1-gado et SWI. Le groupement avec STM-S des séquences SWI est présenté à droite. Les groupes vert et bleu correspondent aux voxels ayant une signature temporelle anormale (ni veine ni lésion). Les encadrements rouges identifient la date d'apparition des anormalités dans les séquences SWI.

De plus, nous pouvons remarquer Fig. 14.2, 14.3 et 14.4 que les comportements spécifiques aux interfaces veines/lésions apparaissent en SWI simultanément à la prise de contraste en T1-gado. Notre hypothèse est que cette hyper intensité transitoire pourrait être interprétée comme une compression de la veine centrale lors de l'apparition de la lésion. Dans la mesure où le rétrécissement des veines intra lésionnelles s'inverse aux points temporels suivant leurs apparitions, quand les lésions sont les plus grandes, il est peu probable que ces hyper intensités transitoires soient des artéfacts dus à des effets de volume partiel.

14.4 Conclusion

Dans ce chapitre, nous avons montré que l'approche STM-S réussit à filtrer et à identifier des groupes de voxels ayant des évolutions d'intensités similaires avec succès. De plus, la faible population des groupes à retrouver par rapport aux autres groupes de l'image rend la tâche difficile et le fait de chercher à extraire un phénomène en limite de résolution spatiale aurait pu mélanger les groupes recherchés avec le reste des images. STM-S prouve donc sa forte capacité à discriminer des évolutions différentes lorsque les paramètres d'échelle sont adaptés. Cette étude suggère que le rétrécissement des veines est un évènement ayant lieu en début de vie des lésions pouvant précéder la rupture de la barrière hémato encéphalique, visible par la prise de contraste en T1-gado. Nous faisons l'hypothèse, via l'étude d'IRM longitudinales de lésions SEP, que la sténose des veines intra lésionnelles est réversible et cohérente avec l'agrégation de globules blancs lors de la formation des lésions SEP. Néanmoins, le rétrécissement transitoire des veines et l'hyper

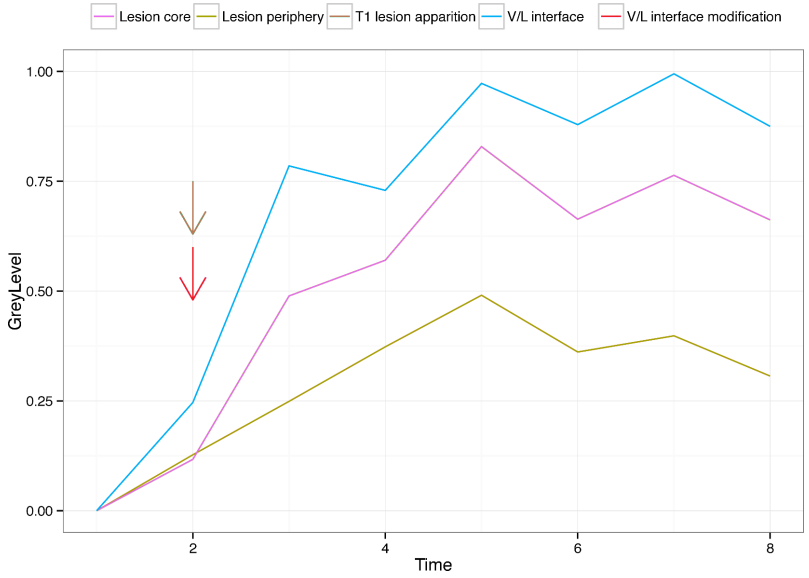


FIGURE 14.2 – Evolution du niveau de gris moyen des classes dans la ROI A. La flèche rouge indique le point temporel de la sténose en SWI et la flèche marron celui de l'apparition de la prise de contraste en T1-gado. Les courbes bleue, violette et jaune correspondent aux classes d'évolution à l'interface veine/lésion, au cœur de la lésion et à sa périphérie.

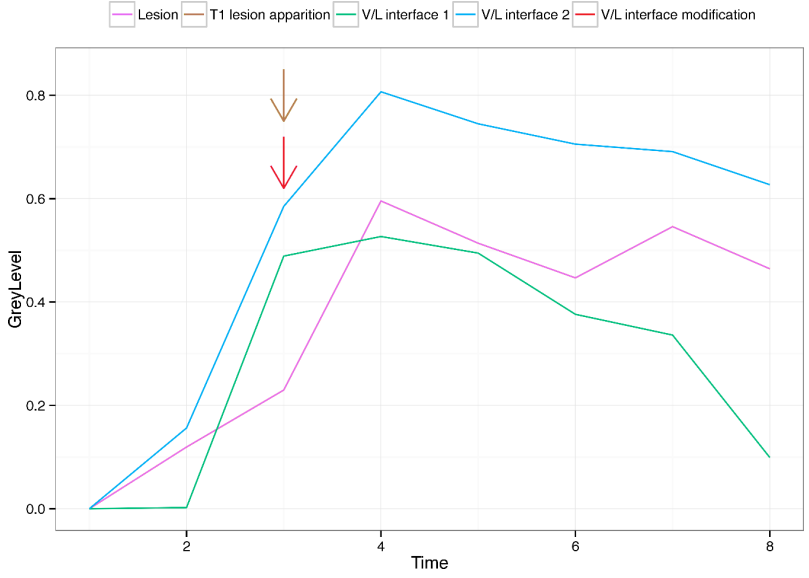


FIGURE 14.3 – Evolution du niveau de gris moyen des classes dans la ROI B. La flèche rouge indique le point temporel de la sténose en SWI et la flèche marron celui de l'apparition de la prise de contraste en T1-gado. Les courbes bleue et verte correspondent aux classes d'évolutions à 2 interfaces veine/lésion, et la courbe violette à la classe d'évolution de la lésion.

intensité intraveineuse observés dans les lésions actives sont cohérents avec les résultats présentés sur des données transversales (Gaitan *et al.* [79]). De plus, la haute fréquence d'échantillonnage temporel comparé à l'évolution de la maladie permet d'observer le développement centrifuge des lésions SEP autour de leur veine centrale, ce qui est aussi

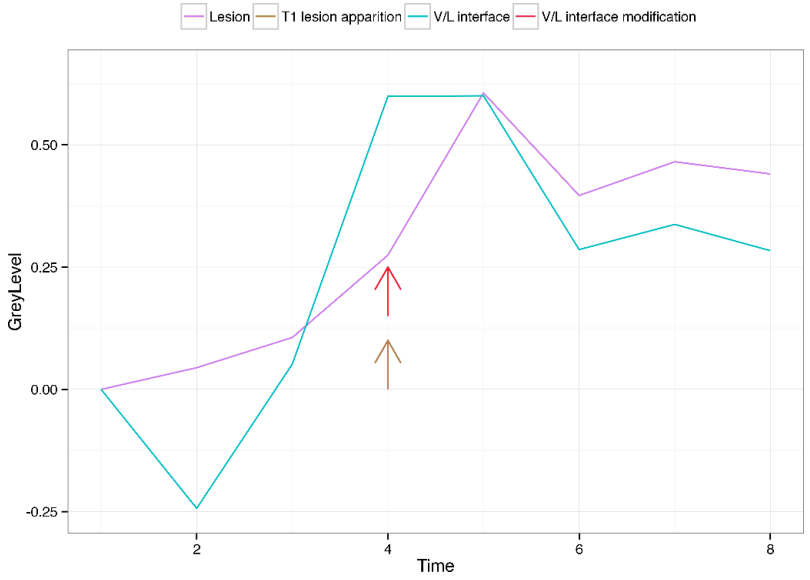


FIGURE 14.4 – Evolution du niveau de gris moyen des classes dans la ROI C. La flèche rouge indique le point temporel de la sténose en SWI et la flèche marron celui de l'apparition de la prise de contraste en T1-gado. Les courbes verte et violette correspondent aux classes d'évolutions à l'interface veine/lésion et au sein de la lésion.

cohérent avec les observations de la littérature (Guttmann *et al.* [72]).

Chapitre 15

Conclusion

Dans ce chapitre, nous avons appliqué et démontré l'utilisabilité de l'approche STM-S et de l'algorithme de groupement basé sur la DTW avec des données réelles de patients atteints par la sclérose en plaques. Nous avons mis en avant des résultats cohérents avec les observations faites dans la littérature mais de manière automatique, sans utiliser d'a priori pour guider la méthode ou sur la finalité du résultat. Ces travaux ont été publiés à la conférence internationale ISBI 2016 [62].

Dans un second temps, nous avons appliqué l'approche STM-S pour l'étude des interactions spécifiques présentes entre les veines cérébrales et les lésions SEP. Nous avons aussi montré qu'il existe un phénomène transitoire entre les veinules cérébrales et les lésions SEP apparaissant aux prémices de leur apparition. Le phénomène suspecté est une sténose des veinules cérébrales qui, s'il s'avérait confirmé dans de prochaines études pourrait aider à comprendre les mécanisme d'apparition des lésions SEP. Ces travaux ont été publiés à la conférence internationale ISMRM 2016 [83] et à la conférence nationale SFNR 2016 [84].

Pour conclure, les méthodes proposées permettent de contribuer à l'analyse automatique des lésions SEP. Les premiers résultats encouragent à traiter un nombre plus important de patients et à varier les types d'entrées afin d'approfondir les travaux sur la compréhension de l'évolution ou des causes de cette maladie.

V Conclusion générale

15.1 Bilan

Ces travaux de thèse se sont focalisés sur l'étude de nouvelles méthodes pour le groupement de séries temporelles issues de séquences d'images acquises au cours du temps.

Nous avons présenté trois nouvelles méthodes permettant de grouper des données temporelles. Notre première contribution, STM-S, permet de grouper des séries temporelles sans a priori sur le nombre de classes à retrouver (Mure *et al.* [54, 60]). Elle nécessite de régler deux paramètres, permettant d'ajuster une tolérance sur les dissimilarités spatiale et temporelle des observations. Les nouveautés apportées par STM-S sont : la modification de la représentation duale spatial/amplitude utilisée par mean-shift pour prendre en considération des données temporelles et l'ajout d'une contrainte de forme des séries temporelles dans le calcul de dissimilarité des évolutions. Cette contrainte est définie par l'introduction de la norme infini dans le formalisme.

STM-S utilisant la distance euclidienne pour comparer les séries temporelles, donc sensible aux modulations temporelles des signaux, nous avons ensuite proposé un algorithme de groupement hiérarchique utilisant la DTW comme critère de groupement des séries temporelles (Mure *et al.* [61]). La nouveauté apportée par cette méthode porte sur le définition d'une contrainte sur la dissimilarité des appariements optimaux trouvés entre deux séquences par la DTW. Comme pour STM-S, nous avons retenu la norme infini pour contraindre la dissimilarité entre les valeurs appariées. L'avantage est qu'il n'est pas nécessaire de spécifier un nombre de groupes pour que l'algorithme hiérarchique s'arrête de fusionner les évolutions, les groupes sont créés par association de proche en proche tant que la condition sur la dissimilarité des séries temporelles est satisfaite.

Les deux méthodes présentées jusqu'alors tiennent uniquement compte de l'ordonnancement temporel des données mais pas des valeurs du temps dans le processus de groupement. Par conséquent, nous avons proposé une méthode, STM-S^{++}, qui redéfinit le formalisme mean-shift pour que les valeurs temporelles des observations soient prises en compte de manière explicite et participent au processus de filtrage (Mure *et al.* [63]). L'avantage d'utiliser les valeurs du temps est que l'on peut définir un critère de similarité prenant en compte l'écart temporel des échantillons, ce qui ne contraint plus les acquisitions à être parfaitement synchronisées.

Nous avons appliqué les méthodes de groupement proposées à l'étude des évolutions des lésions de SEP en suivi longitudinal IRM (Mure *et al.* [62]). Les premiers tests effectués sur ces données sont encourageants et en accord avec les observations de la littérature. Nous avons montré que STM-S retrouve de manière automatique des groupes correspondant aux lésions SEP, dont les évolutions sont similaires à un modèle d'évolution déjà défini [67]. Nous avons aussi utilisé la méthode de groupement proposée section 8 pour mettre en évidence l'existence de similitudes comportementales entre les lésions SEP. Ces similitudes ont été observées entre lésions extraites d'un même patient ou bien entre patients différents.

Enfin, nous avons utilisé STM-S dans le but d'identifier une interaction particulière entre les veinules cérébrales et les lésions SEP (Mure *et al.* [83], Ameli *et al.* [84]). Cette

étude a permis d'identifier des groupes ayant une évolution particulière à l'interface veine/lésion. Ces premières études sont encourageantes et apportent une base méthodologique solide pour conduire des travaux, sur la diversité de comportements des lésions SEP et les facteurs qui pourraient en être à l'origine.

Au delà des applications aux données IRM SEP présentées dans ce manuscrit, nous avons pu appliquer nos contributions à d'autres domaines applicatifs soulignant ainsi la versatilité de nos approches : données satellitaires [60], de spectroscopie [85] et séquences vidéo [63].

15.2 Perspectives

Le point de développement méthodologique le plus naturel suite à ces travaux de thèse serait d'envisager une modification de la mesure de similarité de STM-S afin de pouvoir faire converger directement des séries temporelles déphasées ou modulées localement sur l'axe temporel. L'utilisation d'une mesure telle que DTW dans STM-S n'est qu'une sous partie du problème car le calcul d'une séquence moyenne d'un groupe de séries temporelles sous contrainte d'alignements élastiques n'est pas trivial, certains travaux récents ont apporté des réponses sur le sujet [23, 24].

Néanmoins, l'utilisation d'une métrique ayant une complexité en temps quadratique dans STM-S, qui a aussi une complexité quadratique, pose un problème de temps d'exécution avec des bases de données conséquentes. Un enjeu serait donc l'étude de moyens d'accélération de STM-S pour permettre une convergence plus rapide tout en conservant une estimation convenable de la densité de probabilité. Il est aussi possible de peser sur le calcul de la mesure : par exemple le calcul de la distance euclidienne est suffisant dans certains cas, ce qui apporterait un gain de temps considérable sur le temps d'exécution de la méthode. Cette nécessité d'accélération est indispensable pour le passage à l'échelle sur des bases de données complètes de patients SEP.

Un deuxième axe est l'extension de STM-S^{++} au traitement de séries temporelles. En effet, STM-S^{++} considère les valeurs spatiale, en amplitude et temporelle des observations, mais aucune information sur le lien entre leurs valeurs aux instants t et $t+1$ n'est connue. Il faudrait d'une part, proposer une méthode qui puisse reconstruire ce lien entre les observations à différents instants et d'autre part, décider d'une stratégie de mise à jour des observations vers une valeur moyenne. Dans le cas où les séries temporelles ne sont pas connues en avance, on peut imaginer s'inspirer des méthodes par patch issues des domaines du débruitage de séquences vidéo ou du tracking pour reconstruire les séries temporelles.

Un autre axe d'étude intéressant serait la prise en compte de données hétérogènes pour l'étude de la SEP. Dans les cas où les données ont des caractéristiques suffisamment corrélées, il est simple de les combiner en définissant pour chaque point temporel un vecteur de caractéristiques. Le calcul des distances entre ces vecteurs de caractéristiques est trivial. Néanmoins, si les canaux acquis ne sont pas corrélés, il n'est plus logique de mélanger les données de la sorte. Ainsi, une réflexion autour de méthodes de fusion des

résultats intermédiaires obtenus indépendamment sur les canaux est nécessaire. Cela peut être le cas en IRM, où les séquences acquises peuvent avoir des interprétations anatomiques différentes et ne devraient donc pas être mélangées. Toutefois, elles fournissent chacune des informations utiles sur une pathologie à leur niveau. C'est par exemple le cas dans les bases de données médicales pour l'étude de la SEP.

Concernant la suite des études préliminaires menées sur la sclérose en plaques, les résultats obtenus encouragent à poursuivre avec des bases de données plus importantes que ce que nous avons utilisé jusqu'alors. Il serait intéressant de pouvoir faire une taxonomie des sous-types des lésions SEP grâce à la prise en compte de données multimodales et de comprendre quels sont les facteurs qui sont à l'origine et différencient ces sous-types. Nous pouvons envisager d'utiliser d'autres types d'information que les données images. Leur association permettrait une meilleure discrimination des sous-types de lésions SEP et une compréhension plus approfondie des mécanismes d'évolution de cette pathologie. Beaucoup de questions applicatives restent ouvertes sur la sclérose en plaques. L'utilisation de nouvelles méthodes permettant des caractérisations spatio-temporelles de cette pathologie aux échelles microscopique, lésions et patients pourrait être très intéressante. Par conséquent, les besoins sont doubles sur les aspects applicatifs SEP. D'une part, il faut mener des analyses sur des bases de données plus conséquentes et d'autre part, proposer des méthodes à même de prendre en considération de manière efficace et robuste la diversité d'informations issues de données hétérogènes.

Valorisations scientifiques liées aux travaux

Revue internationale

S. Mure, T. Grenier, D. S. Meier, C.R. Guttmann and H. Benoit-Cattin "Unsupervised spatio-temporal filtering of image sequences : A mean-shift specification." *Pattern Recognition Letters* , vol. 68, Part 1, pp. 48 - 55, 2015

Conférences internationales

S. Mure, T. Grenier, C.R. Guttmann and H. Benoit-Cattin "Unsupervised time-series clustering of distorted and asynchronous temporal patterns." in *2016 IEEE International Conference on ACoustics, Speech and Signal Processing (ICASSP)*, pp. 1263-1267, 2016

S. Mure, T. Grenier, C.R. Guttmann and H. Benoit-Cattin "Unsupervised spatiotemporal video clustering : A versatile mean-shift formulation robust to total object occlusions." in *2016 IEEE International Conference on ACoustics, Speech and Signal Processing (ICASSP)*, pp. 1536-1540, 2016

S. Mure, T. Grenier, C. R. Guttmann, F. Cotton and H. Benoit-Cattin "Classification of multiple sclerosis lesion evolution patterns : A study based on unsupervised clustering of asynchronous time-series." in *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, pp. 1315-1319, 2016

S. Mure, C.R. Guttmann, T. Grenier, H. Benoit-Cattin and F. Cotton "New insight in perivenular lesion formation in MS on weekly susceptibility weighted images." in *24th Annual Meeting of the International Society on Magnetic Resonance in Medicine (ISMRM)*, 2016

A. Dolet, F. Varray, **S. Mure**, T. Grenier, Y. Liu, Z. Yuann, P. Tortoli and D. Vray "Spatial and spectral regularization for multispectral photoacoustic image clustering." in *2016 IEEE International Ultrasonics Symposium (IUS)*, 2016

P. Portejoie, **S. Mure**, H. Benoit-Cattin and T. Grenier "Locally controlled regularized spatiotemporal anisotropic diffusion." *in 2015 IEEE International Conference on Image Processing (ICIP)*, pp. 4823-4827, 2015

Conférences nationales

R. Ameli, **S. Mure**, C.R. Guttmann, T. Grenier, H. Benoit-Cattin and F. Cotton "Analyse dynamique hebdomadaire du développement péri-veinulaire des lésions actives de SEP par imagerie de susceptibilité magnétique" in *43è Congrès de la Société Française de Neuroradiologie (SFNR), Journal of Neuroradiology*, vol. 43, no. 2, pp. 91-93, 2016

S. Mure, T. Grenier, P. Gançarski and H. Benoit-Cattin "Spécification du filtrage mean-shift pour la classification non supervisée de séries temporelles multidimensionnelles." *Colloque GRETSI 2015*, 2015

Bibliographie du manuscrit

[1] Gustavo EAPA Batista, Eamonn J Keogh, Oben Moses Tataw, and Vinícius MA de Souza, "Cid : an efficient complexity-invariant distance for time series," *Data Mining and Knowledge Discovery*, vol. 28, no. 3, pp. 634–669, 2013.

[2] Byoung-Kee Yi and Christos Faloutsos, "Fast time sequence indexing for arbitrary lp norms," VLDB, 2000.

[3] Andrei N Kolmogorov, "On tables of random numbers," *Sankhyā : The Indian Journal of Statistics, Series A*, pp. 369–376, 1963.

[4] Andrei N Kolmogorov, "Three approaches to the quantitative definition ofinformation'," *Problems of information transmission*, vol. 1, no. 1, pp. 1–7, 1965.

[5] Eamonn Keogh, Stefano Lonardi, and Chotirat Ann Ratanamahatana, "Towards parameter-free data mining," in *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2004, pp. 206–215.

[6] Vladimir I Levenshtein, "Binary codes capable of correcting deletions, insertions and reversals," in *Soviet physics doklady*, 1966, vol. 10, p. 707.

[7] Michail Vlachos, Marios Hadjieleftheriou, Dimitrios Gunopulos, and Eamonn Keogh, "Indexing multidimensional time-series," *The VLDB Journal - The International Journal on Very Large Data Bases*, vol. 15, no. 1, pp. 1–20, 2006.

[8] Francois Petitjean, *Dynamic time warping : apports théoriques pour l'analyse de données temporelles : application a la classification de séries temporelles d'images satellites*, Ph.D. thesis, Strasbourg, 2012.

[9] T.H. Cormen, C.E. Leiserson, R.L. Rivest, and S. Clifford, *Introduction to algorithms*, vol. 6, MIT press Cambridge, 2001.

[10] Hiroaki Sakoe and Seibi Chiba, "A dynamic programming approach to continuous speech recognition," in *Proceedings of the seventh international congress on acoustics*, 1971, vol. 3, pp. 65–69.

[11] Hiroaki Sakoe and Seibi Chiba, "Dynamic programming algorithm optimization for spoken word recognition," *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 26, no. 1, pp. 43–49, 1978.

[12] Xiaoyue Wang, Abdullah Mueen, Hui Ding, Goce Trajcevski, Peter Scheuermann, and Eamonn Keogh, "Experimental comparison of representation methods and distance

measures for time series data," *Data Mining and Knowledge Discovery*, vol. 26, no. 2, pp. 275–309, 2013.

[13] Nurjahan Begum, Liudmila Ulanova, Jun Wang, and Eamonn Keogh, "Accelerating dynamic time warping clustering with a novel admissible pruning strategy," in *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2015, pp. 49–58.

[14] Fumitada Itakura, "Minimum prediction residual principle applied to speech recognition," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 23, no. 1, pp. 67–72, 1975.

[15] Ahlame Douzal Chouakria and Panduranga Naidu Nagabhushan, "Adaptive dissimilarity index for measuring time series proximity," *Advances in Data Analysis and Classification*, vol. 1, no. 1, pp. 5–21, 2007.

[16] Ahlame Douzal-Chouakria, *Contribution à l'analyse de données temporelles*, Ph.D. thesis, Université Joseph-Fourier-Grenoble I, 2012.

[17] Jaewon Yang and Jure Leskovec, "Patterns of temporal variation in online media," in *Proceedings of the fourth ACM international conference on Web search and data mining*. ACM, 2011, pp. 177–186.

[18] John Paparrizos and Luis Gravano, "k-shape : Efficient and accurate clustering of time series," in *Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data*. ACM, 2015, pp. 1855–1870.

[19] T Warren Liao, "Clustering of time series data - a survey," *Pattern recognition*, vol. 38, no. 11, pp. 1857–1874, 2005.

[20] Chung Laung Liu, "Introduction to combinatorial mathematics," 1968.

[21] Leonard Kaufman and Peter J Rousseeuw, "Partitioning around medoids (program pam)," *Finding groups in data : an introduction to cluster analysis*, pp. 68–125, 1990.

[22] Rui Xu and DC Wunsch, "Clustering. hoboken," 2009.

[23] François Petitjean, Alain Ketterlin, and Pierre Gançarski, "A global averaging method for dynamic time warping, with applications to clustering," *Pattern Recognition*, vol. 44, no. 3, pp. 678–693, 2011.

[24] Saeid Soheily-Khah, Ahlame Douzal-Chouakria, and Eric Gaussier, "Generalized k-means-based clustering for temporal data under weighted and kernel time warp," *Pattern Recognition Letters*, vol. 75, pp. 63–69, 2016.

[25] Frank Harary et al., "Graph theory," 1969.

[26] Anil K Jain and Richard C Dubes, *Algorithms for clustering data*, Prentice-Hall, Inc., 1988.

[27] Andrew Y Ng, Michael I Jordan, Yair Weiss, et al., "On spectral clustering : Analysis and an algorithm," *Advances in neural information processing systems*, vol. 2, pp. 849–856, 2002.

[28] Boaz Nadler and Meirav Galun, "Fundamental limitations of spectral clustering," in *Advances in neural information processing systems*, 2006, pp. 1017–1024.

[29] Martin Ester, Hans-Peter Kriegel, Jörg Sander, Xiaowei Xu, et al., "A density-based algorithm for discovering clusters in large spatial databases with noise.," in *Kdd*, 1996, vol. 96, pp. 226–231.

[30] Mihael Ankerst, Markus M. Breunig, Hans peter Kriegel, and Jorg Sander, "Optics : Ordering points to identify the clustering structure," 1999, pp. 49–60, ACM Press.

[31] Levent Ertöz, Michael Steinbach, and Vipin Kumar, "Finding clusters of different sizes, shapes, and densities in noisy, high dimensional data.," in *SDM*. SIAM, 2003, pp. 47–58.

[32] C. Aggarwal and C. Reddy, *Data Clustering : Algorithms and Applications*, CRC Press, 2013.

[33] Panos Kalnis, Nikos Mamoulis, and Spiridon Bakiras, "On discovering moving clusters in spatio-temporal data," in *International Symposium on Spatial and Temporal Databases*. Springer, 2005, pp. 364–381.

[34] Hoyoung Jeung, Man Lung Yiu, Xiaofang Zhou, Christian S Jensen, and Heng Tao Shen, "Discovery of convoys in trajectory databases," *Proceedings of the Very Large Data Bases Endowment*, vol. 1, no. 1, pp. 1068–1080, 2008.

[35] Marc Benkert, Joachim Gudmundsson, Florian Hübner, and Thomas Wolle, "Reporting flock patterns," *Computational Geometry*, vol. 41, no. 3, pp. 111–125, 2008.

[36] Zhenhui Li, Bolin Ding, Jiawei Han, and Roland Kays, "Swarm : Mining relaxed temporal moving object clusters," *Proceedings of the Very Large Data Bases Endowment*, vol. 3, no. 1-2, pp. 723–734, 2010.

[37] Kai Zheng, Yu Zheng, Nicholas Jing Yuan, and Shuo Shang, "On discovery of gathering patterns from trajectories," in *Data Engineering, IEEE Proceedings of the 29th International Conference on*, 2013, pp. 242–253.

[38] Alex Rodriguez and Alessandro Laio, "Clustering by fast search and find of density peaks," *Science*, vol. 344, no. 6191, pp. 1492–1496, 2014.

[39] K. Fukunaga and L. D. Hostetler, "Estimation of the gradient of a density function with applications in pattern recognition.," *Information Theory, IEEE Transaction on*, vol. 21, no. 1, pp. 32–40, 1975.

[40] Y. Cheng, "Mean shift, mode seeking, and clustering," *Pattern Analysis and Machine Intelligence, IEEE Transaction on*, vol. 17, no. 8, pp. 790–799, 1995.

[41] D. Comaniciu and P. Meer, "Mean shift : A robust approach toward feature space analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 603–619, 2002.

[42] Mark Fashing and Carlo Tomasi, "Mean shift is a bound optimization," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 27, no. 3, pp. 471–474, 2005.

[43] T. Grenier, C. Revol-Muller, and G. Gimenez, "Hybrid approach for multiparametric mean shift filtering," in *Image Processing, IEEE Proceedings of the International Conference on*, Atlanta, USA, 2006, pp. 1541–1544.

[44] Dorin Comaniciu, Visvanathan Ramesh, and Peter Meer, "The variable bandwidth mean shift and data-driven scale selection," in *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*. IEEE, 2001, vol. 1, pp. 438–445.

[45] D. Comaniciu, V. Ramesh, and P. Meer, "Kernel-based object tracking," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 25, no. 5, pp. 564–577, May 2003.

[46] Jue Wang, Bo Thiesson, Yingqing Xu, and Michael Cohen, "Image and video segmentation by anisotropic kernel mean shift," in *Computer Vision-ECCV 2004*, pp. 238–249. Springer, 2004.

[47] Thomas Grenier, Chantal Revol-Muller, Franck Davignon, Olivier Basset, and Gérard Gimenez, "Variable bandwidth mean shift for smoothing ultrasonic images," in *Proceedings of the EUSIPCO*, 2005, vol. 5.

[48] Ting Li, Sorina Camarasu-Pop, Tristan Glatard, Thomas Grenier, and Hugues Benoit-Cattin, "Optimization of mean-shift scale parameters on the egee grid.," in *Health-Grid*, 2010, pp. 203–214.

[49] Wei Feng and Rong-Chun Zhao, "Non-rigid objects detection and segmentation in video sequence using 3d mean shift analysis," in *International Conference on Machine Learning and Cybernetics*, Nov 2003, vol. 5, pp. 3134–3139 Vol.5.

[50] IreneY.H. Gu, Vasile Gui, and Zhifei Xu, "Video segmentation using joint space-time-range adaptive mean shift," in *Advances in Multimedia Information Processing - PCM 2006*, Yueting Zhuang, Shi-Qiang Yang, Yong Rui, and Qinming He, Eds., vol. 4261 of *Lecture Notes in Computer Science*, pp. 740–748. Springer Berlin Heidelberg, 2006.

[51] Sylvain Paris, "Edge-preserving smoothing and mean-shift segmentation of video streams," in *Computer Vision - ECCV 2008*, David Forsyth, Philip Torr, and Andrew Zisserman, Eds., vol. 5303 of *Lecture Notes in Computer Science*, pp. 460–473. Springer Berlin Heidelberg, 2008.

[52] M. Grundmann, V. Kwatra, Mei Han, and I. Essa, "Efficient hierarchical graph-based video segmentation," in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, June 2010, pp. 2141–2148.

[53] Tomoyuki Nagahashi, Hironobu Fujiyoshi, and Takeo Kanade, "Video segmentation using iterated graph cuts based on spatio-temporal volumes," in *Computer Vision - ACCV 2009*, Hongbin Zha, Rin-ichiro Taniguchi, and Stephen Maybank, Eds., vol. 5995 of *Lecture Notes in Computer Science*, pp. 655–666. Springer Berlin Heidelberg, 2010.

[54] Simon Mure, Thomas Grenier, Dominik S. Meier, Charles R.G. Guttmann, and Hugues Benoit-Cattin, "Unsupervised spatio-temporal filtering of image sequences. a mean-shift specification," *Pattern Recognition Letters*, vol. 68, Part 1, pp. 48 – 55, 2015.

[55] K.K. Leung, N. Saeed, K. Changani, S.P. Campbell, and D.L.G. Hill, "Spatio-temporal segmentation of rheumatoid arthritis lesions in serial mr images of joints," in *Computer Vision and Pattern Recognition Workshop, 2006. IEEE Conference on*, June 2006, pp. 91–91.

[56] Jian Cheng, Feng Shi, Kun Wang, Ming Song, Jiefeng Jiang, Lijuan Xu, and Tianzi Jiang, "Nonparametric mean shift functional detection in the functional space for task and resting-state fmri," in *Workshop on fMRI data analysis : statistical modeling and detection issues in intra- and inter-subject functional MRI data analysis, in conjunction with the MICCAI 2009*, London, United Kingdom, Sept. 2009.

[57] Leo Ai, Xin Gao, and Jinhu Xiong, "Application of mean-shift clustering to blood oxygen level dependent functional mri activation detection," *BMC Medical Imaging*, vol. 14, no. 1, pp. 6, 2014.

[58] L Shepp and B. F. Logan, "The fourier reconstruction of a head section," *Nuclear Science, IEEE Transactions on*, vol. 21, pp. 21–43, 1974.

[59] T Li, T. Grenier, and H. Benoit-Cattin, "Color space influence on mean shift filtering," in *Image Processing, IEEE Proceedings of the International Conference on*, Brussels, Belgium, 2011, pp. 1469 – 1472.

[60] S. Mure, T. Grenier, P. Gançarski, and H. Benoit-Cattin, "Spécification du filtrage mean-shift pour la classification non supervisée de séries temporelles multidimensionnelles," in *Colloque GRETSI 2015*, 2015.

[61] S. Mure, T. Grenier, C. R. G. Guttmann, and H. Benoit-Cattin, "Unsupervised time-series clustering of distorted and asynchronous temporal patterns," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2016, pp. 1263–1267.

[62] S. Mure, T. Grenier, C. R. G. Guttmann, F. Cotton, and H. Benoit-Cattin, "Classification of multiple sclerosis lesion evolution patterns : A study based on unsupervised clustering of asynchronous time-series," in *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, 2016, pp. 1315–1319.

[63] S. Mure, T. Grenier, and H. Benoit-Cattin, "Unsupervised spatiotemporal video clustering : A versatile mean-shift formulation robust to total object occlusions," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2016, pp. 1536–1540.

[64] Kristen Coyne, "Mri : A guided tour," *Internet : http ://www. magnet. fsu. edu/education/tutorials/magnetacademy/mri/*, 2012.

[65] Alastair Compston and Alasdair Coles, "Multiple sclerosis," *The Lancet*, vol. 372, no. 9648, pp. 1502–1517, 2008.

[66] Lars Bø, Christian A Vedeler, Harald I Nyland, Bruce D Trapp, and Sverre J Mørk, "Subpial demyelination in the cerebral cortex of multiple sclerosis patients," *Journal of Neuropathology & Experimental Neurology*, vol. 62, no. 7, pp. 723–732, 2003.

[67] Dominik S Meier, HL Weiner, and Charles RG Guttmann, "Mr imaging intensity modeling of damage and repair in multiple sclerosis : relationship of short-term lesion recovery to progression and disability," *American Journal of Neuroradiology*, vol. 28, no. 10, pp. 1956–1963, 2007.

[68] Dominik S Meier and Charles RG Guttmann, "Time-series analysis of mri intensity patterns in multiple sclerosis," *NeuroImage*, vol. 20, no. 2, pp. 1193–1209, 2003.

[69] Guido Gerig, Daniel Welti, Charles RG Guttmann, Alan CF Colchester, and Gábor Székely, "Exploring the discrimination power of the time domain for segmentation and characterization of active lesions in serial mr data," *Medical Image Analysis*, vol. 4, no. 1, pp. 31–42, 2000.

[70] Howard L Weiner, Charles RG Guttmann, Samia J Khoury, E John Orav, Marika J Hohol, Ron Kikinis, and Ferenc A Jolesz, "Serial magnetic resonance imaging in multiple sclerosis : correlation with attacks, disability, and disease stage," *Journal of neuroimmunology*, vol. 104, no. 2, pp. 164–173, 2000.

[71] DS Meier, KE Balashov, B Healy, HL Weiner, and CRG Guttmann, "Seasonal prevalence of ms disease activity," *Neurology*, vol. 75, no. 9, pp. 799–806, 2010.

[72] CR Guttmann, Sungkee S Ahn, Liangge Hsu, Ron Kikinis, and Ferenc A Jolesz, "The evolution of multiple sclerosis lesions on serial mr.," *American journal of neuroradiology*, vol. 16, no. 7, pp. 1481–1491, 1995.

[73] Stephen M Smith, "Fast robust automated brain extraction," *Human brain mapping*, vol. 17, no. 3, pp. 143–155, 2002.

[74] Mark Jenkinson and Stephen Smith, "A global optimisation method for robust affine registration of brain images," *Medical image analysis*, vol. 5, no. 2, pp. 143–156, 2001.

[75] Mark Jenkinson, Peter Bannister, Michael Brady, and Stephen Smith, "Improved optimization for the robust and accurate linear registration and motion correction of brain images," *Neuroimage*, vol. 17, no. 2, pp. 825–841, 2002.

[76] Yulin Ge, Meng Law, Joseph Herbert, and Robert I Grossman, "Prominent perivenular spaces in multiple sclerosis as a sign of perivascular inflammation in primary demyelination," *American journal of neuroradiology*, vol. 26, no. 9, pp. 2316–2319, 2005.

[77] JW Dawson, "The histology of multiple sclerosis," *Trans R Soc Edinburgh*, vol. 50, pp. 517–740, 1916.

[78] Douglas McAlpine, Nigel D Compston, and Charles E Lumsden, "Multiple sclerosis, london, e. & s," *Livingstone, lad*, vol. 195, 1955.

[79] María I Gaitán, Manori P de Alwis, Pascal Sati, Govind Nair, and Daniel S Reich, "Multiple sclerosis shrinks intralesional, and enlarges extralesional, brain parenchymal veins," *Neurology*, vol. 80, no. 2, pp. 145–151, 2013.

[80] E Mark Haacke, Yingbiao Xu, Yu-Chung N Cheng, and Jürgen R Reichenbach, "Susceptibility weighted imaging (swi)," *Magnetic resonance in medicine*, vol. 52, no. 3, pp. 612–618, 2004.

[81] Francois Cotton, Howard L Weiner, Ferenc A Jolesz, and Charles RG Guttmann, "Mri contrast uptake in new lesions in relapsing-remitting ms followed at weekly intervals," *Neurology*, vol. 60, no. 4, pp. 640–646, 2003.

[82] Charles RG Guttmann, Matthieu Rousset, Jean A Roch, Salem Hannoun, Françoise Durand-Dubief, Boubakeur Belaroussi, Michele Cavallari, Muriel Rabilloud, Dominique Sappey-Marinier, Sandra Vukusic, et al., "Multiple sclerosis lesion formation and early evolution revisited : A weekly high-resolution magnetic resonance imaging study," *Multiple Sclerosis Journal*, vol. 22, no. 6, pp. 761–769, 2016.

[83] S. Mure, C. R. G. Guttmann, T. Grenier, H. Benoit-Cattin, and F. Cotton, "New insight in perivenular lesion formation in ms on weekly susceptibility weighted images," in *24th Annual Meeting International Society on Magnetic Resonance in Medicine (ISMRM)*, 2016.

[84] R. Ameli, S. Mure, CRG. Guttmann, T. Grenier, H. Benoit-Cattin, and F. Cotton, "Analyse dynamique hebdomadaire du développement péri-veinulaire des lésions actives de sep par imagerie de susceptibilité magnétique," *Journal of Neuroradiology*, vol. 43, no. 2, pp. 91–93, 2016.

[85] A. Dolet, F. Varray, S. Mure, T. Grenier, Y. Liu, Z. Yuan, P. Tortoli, and D. Vray, "Spatial and spectral regularization for multispectral photoacoustic image clustering," in *2016 IEEE International Ultrasonics Symposium (IUS)*, 2016.

FOLIO ADMINISTRATIF

THESE DE L'UNIVERSITE DE LYON OPEREE AU SEIN DE L'INSA LYON

NOM : MURE

DATE de SOUTENANCE : 2 décembre 2016

Prénom : Simon

TITRE : Classification non supervisée de données spatio-temporelles multidimensionnelles. Applications à l'imagerie.

NATURE : Doctorat

Numéro d'ordre : 2016LYSEI130

Ecole doctorale : Électronique, Électrotechnique, Automatique

Spécialité : Traitement du signal et de l'image

RESUME : Avec l'augmentation considérable d'acquisitions de données temporelles dans les dernières décennies comme les systèmes GPS, les séquences vidéo ou les suivis médicaux de pathologies ; le besoin en algorithmes de traitement et d'analyse efficaces d'acquisition longitudinales n'a fait qu'augmenter. Dans cette thèse, nous proposons une extension du formalisme mean-shift, classiquement utilisé en traitement d'images, pour le groupement de séries temporelles multidimensionnelles. Nous proposons aussi un algorithme de groupement hiérarchique des séries temporelles basé sur la mesure de dynamic time warping afin de prendre en compte les déphasages temporels. Ces choix ont été motivés par la nécessité d'analyser des images acquises en imagerie par résonance magnétique sur des patients atteints de sclérose en plaques. Cette maladie est encore très méconnue tant dans sa genèse que sur les causes des handicaps qu'elle peut induire. De plus aucun traitement efficace n'est connu à l'heure actuelle. Le besoin de valider des hypothèses sur les lésions de sclérose en plaque nous a conduit à proposer des méthodes de groupement de séries temporelles ne nécessitant pas d'a priori sur le résultat final, méthodes encore peu développées en traitement d'images.

MOTS-CLÉS : Séries temporelles, classification non supervisée, mean-shift, dynamic time warping, norme infini, sclérose en plaques, imagerie par résonance magnétique.

Laboratoire (s) de recherche : CREATIS

Directeurs de thèse : Pr. Hugues BENOIT-CATTIN et Dr. Thomas Grenier

Président de jury : Pr. Christophe DUCOTTET

Composition du jury :

Christophe DUCOTTET	Professeur des Universités	LHC, Saint Etienne	Président
Ahlame DOUZAL	Maître de conférences, HDR	LIG-AMA, Grenoble	Rapporteure
Ludovic MACAIRE	Professeur des Universités	LAGIS, Lille	Rapporteur
Pierre-François MARTEAU	Professeur des Universités	IRISA, Vannes	Examinateur
Charles GUTTMANN	Associate professor	Harvard Medical School, Boston	Membre invité
Thomas GRENIER	Maître de conférences	CREATIS, INSA Lyon	Co-directeur de thèse
Hugues BENOIT-CATTIN	Professeur des Universités	CREATIS, INSA Lyon	Directeur de thèse