



HAL
open science

Identification et caractérisation du thrombus veineux par imagerie échographique mode B couplée à l'élastographie

Thibaud Berthomier

► **To cite this version:**

Thibaud Berthomier. Identification et caractérisation du thrombus veineux par imagerie échographique mode B couplée à l'élastographie. Traitement du signal et de l'image [eess.SP]. ENSTA Bretagne - École nationale supérieure de techniques avancées Bretagne, 2018. Français. NNT : 2018ENTA0007 . tel-02366284

HAL Id: tel-02366284

<https://theses.hal.science/tel-02366284>

Submitted on 15 Nov 2019

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

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

THESE DE DOCTORAT DE

L'ECOLE NATIONALE SUPERIEURE
DE TECHNIQUES AVANCEES BRETAGNE
COMUE UNIVERSITE BRETAGNE LOIRE

ECOLE DOCTORALE N° 601
*Mathématiques et Sciences et Technologies
de l'Information et de la Communication*
Spécialité : *Signal, Image, Vision*

Par

Thibaud BERTHOMIER

Identification et caractérisation du thrombus veineux par imagerie échographique mode B couplée à l'élastographie

Thèse présentée et soutenue à Brest, le 13 novembre 2018

Unité de recherche : UFR Médecine et Sciences de la Santé – Unité INSERM CIC 1412 – EA 3878 GETBO
INNOVTE – F-CRIN Network – Unité INSERM CIC 1408 – CHU Saint-Etienne

Composition du Jury :

Président du jury :

Christian JUTTEN Professeur, Université Joseph Fourier,
Membre Senior de l'Institut Universitaire de France

Rapporteurs :

Denis HAMAD Professeur, Université du Littoral Côte d'Opale
Marie-Antoinette SEVESTRE Professeur, EA 4666, CHU Amiens-Picardie

Examineurs :

Emanuel RADOI Professeur, Université Bretagne Loire
Hélène THOMAS Enseignant-Chercheur, ENSTA-Bretagne
Benoit ZERR Enseignant-Chercheur, HDR, ENSTA-Bretagne

Co-directeurs de thèse :

Luc BRESSOLLETTE Professeur, EA 3878 CHU de Brest
Ali MANSOUR Professeur, ENSTA-Bretagne

Avant propos

Cette thèse est une contribution à l'analyse des échographies et élastographies dans le but d'identifier et de caractériser la structure d'un thrombus veineux profond. Elle a été réalisée en collaboration entre le Groupe d'Etude de Thrombose de Bretagne Occidentale (EA 3878 GETBO, CIC Inserm 1412) , le réseau INNOVTE (INvestigation Network On Venous Thrombo-Embolism, CIC Inserm 1408) et l'Ecole Nationale Supérieure de Techniques Avancées Bretagne (ENSTA Bretagne, Lab-STICC UMR CNRS 6285).

La thèse a été dirigée par Messieurs Ali Mansour, Professeur de l'ENSTA Bretagne et Luc Bressollette, Professeur de la Faculté de Médecine et Sciences de la Santé (Université Bretagne Loire). Elle a été encadrée par Messieurs Dominique Mottier, Professeur du CHRU de Brest et Frédéric Le Roy, Enseignant-Chercheur de l'ENSTA-Bretagne.

Je tiens tous d'abord à remercier le Centre D'Investigation Clinique (CIC) de Brest, le réseau INNOVTE (INvestigation Network On Venous Thrombo-Embolism) labellisé F-CRIN (French Clinical Research Infrastructure Network) et la Région Bretagne d'avoir rendu possible et financer cette thèse. Je remercie aussi le groupe Canon Medical Systems (anciennement Toshiba Medical Systems) pour le soutien matériel du projet et la constante évolution de leur système d'acquisition.

Je tiens également à remercier Luc Bressollette et Dominique Mottier, qui sont à l'initiative de ce projet, de leur accueil au sein du CHU de Brest, de leur aide et de leur confiance. Je remercie également Ali Mansour et Frédéric Le Roy de m'avoir proposé ce sujet de thèse et de m'accompagner quotidiennement dans mes travaux de thèse.

J'exprime ma gratitude à Monsieur Denis Hamad, Professeur de l'Université du Littoral Côte d'Opale (ULCO), et à Madame Marie-Antoinette Sevestre, Professeur du CHU Amiens-Picardie, pour m'avoir fait le privilège d'être les rapporteurs de cette thèse, ainsi qu'à Christian Jutten, Professeur de l'Université Joseph Fourier, Emanuel Radoi, Professeur de l'Université Bretagne Loire, Benoit Zerr, Professeur de l'ENSTA Bretagne et Hélène Thomas, Enseignant-Chercheur de l'ENSTA Bretagne, d'avoir accepté d'examiner ce travail.

Je remercie également le Professeur Sandrine Jousse-Joulin, les Docteurs Barthélémy Hermenault, Clément Hoffmann, Léo Fréchier qui m'ont donné de la matière pour mes expérimentations et avec qui ce fut un plaisir de collaborer.



Table des matières

Table des figures	xi
Liste des tableaux	xiii
Glossaire médical	xiv
Glossaire du traitement d'images	xvi
Introduction	1
1 Contexte et problématiques	3
1.1 Thrombose veineuse profonde	4
1.1.1 Formation adaptée d'un thrombus veineux	4
1.1.2 Formation inadaptée	4
1.1.2.1 La stase sanguine	4
1.1.2.2 L'altération des parois veineuses	5
1.1.2.3 La thrombophilie	5
1.1.3 Observation clinique et risques évolutifs	6
1.1.4 Diagnostic et traitements	7
1.1.5 Conclusion sur la problématique médicale de la thèse	8
1.2 Imagerie acoustique pour la caractérisation de tissus humains	9
1.2.1 Propagation acoustique	9
1.2.1.1 En milieu élastique	9
1.2.1.2 Dans le corps humain	10
1.2.2 Echographie	11
1.2.2.1 Principe	11
1.2.2.2 Imageries fondamentales et harmoniques	12
1.2.3 Elastographie	13
1.2.3.1 Méthodes quasi-statiques	13
1.2.3.2 Méthodes dynamiques ou impulsionnelles	14
1.2.4 Conclusion sur l'imagerie acoustique en médecine	15
1.3 Collecte des images échographiques et élastographiques	16
1.3.1 Système d'acquisition	16
1.3.1.1 Echographe Toshiba Aplio 500	16
1.3.1.2 Sondes	18
1.3.1.3 Paramètres d'acquisition	19
1.3.2 Evolution du protocole d'acquisition des données de thrombi	21
1.3.2.1 Protocole 1 : pré-mise à jour	21
1.3.2.2 Protocole 2 : post-mise à jour	22
1.3.2.3 Protocole 3 : sonde convexe	24
1.3.3 Acquisition des images de glandes salivaires	26
1.3.4 Conclusion sur les bases de données	28
1.4 Conclusion partielle sur la partie médicale de notre projet	28

2	Etat de l'art	29
2.1	Analyse du thrombus veineux	30
2.1.1	Par échographie	30
2.1.2	Par élastographie	31
2.1.3	Conclusion sur l'analyse des thrombi veineux	33
2.2	Méthodes de traitement d'images dans le domaine du biomédical	34
2.2.1	Domaines d'application	34
2.2.2	Prétraitements	35
2.2.3	Segmentation	36
2.2.4	Extraction de descripteurs	37
2.2.5	Classification	38
2.2.6	Conclusion sur l'état de l'art en imagerie biomédicale	39
2.3	Réduction de la dimension des descripteurs	40
2.3.1	Réduction par passage dans le domaine de Fourier (ou des cosinus)	40
2.3.2	Analyse en Composante Principale (ACP)	41
2.3.3	T-distributed stochastic neighbor embedding	42
2.3.4	Autres techniques	43
2.3.5	Conclusion sur les techniques de réduction de dimension	43
2.4	Classifieurs	44
2.4.1	Classification supervisée	44
2.4.1.1	Classification par SMA	44
2.4.1.2	Support Vector Machine (SVM)	45
2.4.1.3	Apprentissage profond	46
2.4.2	Classification non supervisée	49
2.4.2.1	Partitionnement par k -moyennes	49
2.4.2.2	Partitionnement par DBSCAN	50
2.4.2.3	Classification spectrale	51
2.4.3	Conclusion sur les méthodes de classification	55
2.5	Conclusion sur l'état de l'art et le choix des approches développées	55
3	Base de données et prétraitements	57
3.1	Description de la base de données	58
3.1.1	Base de thrombus	58
3.1.2	Base de glandes salivaires	60
3.2	Préparation des données	61
3.2.1	Cas des thrombi	61
3.2.1.1	Lecture et conversion	61
3.2.1.2	Découpage, résolution et type d'images	62
3.2.2	Cas des glandes salivaires	63
3.2.3	Egalisation d'histogramme des échographies	64
3.2.4	Conversion des élastographies en cartes de vitesses	64
3.2.5	Conclusion sur la préparation des données	65
3.3	Segmentation et extraction d'images de la région d'intérêt	66
3.3.1	Echographies des thrombi	66
3.3.1.1	Détection et position de la ROI	66
3.3.1.2	Segmentation de la ROI	66
3.3.1.3	Extraction d'images	67
3.3.2	Echographies des glandes salivaires	68
3.3.3	Elastographies des thrombi et des glandes salivaires	69
3.3.4	Conclusion sur la segmentation et l'extraction des images	69
3.4	Conclusion sur la préparation des bases de données	70

4	Caractérisation des thrombi par approches multirésolutions et statistiques	71
4.1	Approche multirésolution : le <i>Scattering Operator</i> (SO)	72
4.1.1	Principe	72
4.1.2	Ondelettes et propriétés du SO	73
4.1.3	Représentation des transformées du SO	76
4.1.4	Invariances en rotation et en échelle	77
4.1.4.1	Invariance en rotation :	77
4.1.4.2	Invariance en échelle :	78
4.1.5	Conclusion sur le <i>Scattering Operator</i> (SO)	79
4.2	Approches par statistiques	80
4.2.1	Moments et cumulants spatiaux	80
4.2.2	Autocorrélation, bicorrélation et tricorrélation	81
4.2.3	Construction des descripteurs	83
4.2.4	Conclusion sur l'approche par statistiques	85
4.3	Combinaison du SO et des Statistiques d'Ordre Supérieur (SOS)	86
4.3.1	Structure des descripteurs	86
4.3.2	Uniformisation des descripteurs	87
4.3.3	Conclusion sur la combinaison des deux approches	88
4.4	Conclusion sur l'extraction de descripteurs multirésolutions et statistiques	88
5	Simulations	89
5.1	Validation et optimisation sur une base de référence	90
5.1.1	Intérêt et présentation de la base KTH-TIPS	90
5.1.2	SO	91
5.1.2.1	Sélection des paramètres clés	91
5.1.2.2	Logarithme, invariances et apprentissage multi-échelle	94
5.1.2.3	Classification non supervisée et paramètres du SO	96
5.1.2.4	Bilan sur les descripteurs multirésolutions	97
5.1.3	Multicorrélations	98
5.1.3.1	Ordre maximal	98
5.1.3.2	Choix des combinaisons d'axes	101
5.1.3.3	Bilan sur les descripteurs statistiques	103
5.1.4	Descripteurs multirésolutions et statistiques	104
5.1.4.1	Uniformisation des descripteurs	104
5.1.4.2	Evaluation des performances	105
5.1.4.3	Bilan sur l'utilisation conjointe du SO et SOS	107
5.1.5	Bilan des simulations sur la base de référence	107
5.2	Evaluation sur les glandes salivaires	108
5.2.1	Diagnostic des glandes salivaires par échographie et élastographie	108
5.2.2	Protocole expérimental	108
5.2.2.1	Rappel sur la base	109
5.2.2.2	Définition des descripteurs	110
5.2.2.3	Uniformisation des descripteurs	110
5.2.2.4	Fusion des descripteurs	111
5.2.2.5	Classification des descripteurs	111
5.2.2.6	Résumé du protocole	112
5.2.3	Résultats du protocole	113
5.2.3.1	Parotides et submandibulaires	113
5.2.3.2	Parotides	119
5.2.3.3	Submandibulaires	121
5.2.4	Bilan des simulations sur les glandes salivaires	123
5.3	Classification du thrombus veineux	124
5.3.1	Base d'étude et protocole expérimental	124
5.3.2	Facteurs déclencheurs des TVP	126

5.3.3	Caractère provoqué de la TVP et présence d'EP	129
5.3.4	Caractère provoqué de la TVP	132
5.3.5	Présence d'EP	134
5.3.6	Bilan des simulations sur les thrombi veineux	136
5.4	Conclusion sur la partie simulation	137
Conclusion générale et perspectives		139
Liste des publications		145
A Thrombus veineux et Imagerie		149
A.1	Hémostase et fibrinolyse	149
A.2	Principaux outils de détection d'un thrombus	150
A.2.1	Phlébographie	150
A.2.2	Doppler veineux	150
A.2.3	Angiographie par Résonance Magnétique	151
A.2.4	Angioscanner	151
A.3	Propagation acoustique dans un milieu inhomogène	151
B Evolution de la chaîne de prétraitements		153
B.1	Lecture et conversion des fichiers DICOM	153
B.2	Suppression des annotations et de l'ellipse	155
B.3	Extraction d'une imagerie à l'intérieur de l'ellipse	155
B.3.1	Segmentation des ellipses	156
B.3.1.1	Cas 1 : contour de l'ellipse non fermée	156
B.3.1.2	Cas 2 : ellipse d'aire insuffisante	157
B.3.1.3	Cas 3 : absence d'ellipse	157
B.3.1.4	Cas 4 : plusieurs ellipses	158
B.3.2	Extraction de l'imagerie	158
B.4	Uniformisation de la taille des images	159
C Exploration multirésolution des signaux		161
C.1	Cas des signaux monodimensionnels	161
C.1.1	Analyse de Fourier	161
C.1.2	Analyse temps-fréquence	163
C.1.3	Analyse temps-échelle	163
C.1.4	Analyse multirésolution	165
C.1.5	Ondelettes et bancs de filtres	167
C.1.6	Algorithme de Mallat	167
C.1.7	Généralisation à M -bandes	168
C.1.8	Paquet d'ondelettes	169
C.2	Cas des signaux bidimensionnels	170
C.2.1	Transformée en ondelettes bidimensionnelles séparables	170
C.2.2	Analyse en ondelettes non séparables	173
D Complément sur les statistiques d'ordre supérieur		175
D.1	Introduction aux statistiques d'ordre supérieur	175
D.1.1	Définitions préliminaires	175
D.1.1.1	Évènement et expérience aléatoire	175
D.1.1.2	Probabilité d'un évènement	176
D.1.1.3	Variable aléatoire et fonction de répartition	176
D.1.1.4	Espérance mathématique	177
D.1.1.5	Variance et moment d'ordre deux	177
D.1.1.6	Loi normale ou gaussienne	178
D.1.2	SOS	179

D.1.2.1	Moments d'ordre supérieur	179
D.1.2.2	Fonctions caractéristiques	179
D.1.2.3	Cumulants d'ordre supérieur	180
D.1.2.4	Propriétés entre moments et cumulants	180
D.1.2.5	Interprétation des cumulants	181
D.1.3	Couple de variables aléatoires	182
D.1.3.1	Lois de probabilité conjointes et marginales	182
D.1.3.2	Espérance et moments d'ordre deux	183
D.1.3.3	Fonctions caractéristiques et cumulants	184
D.1.4	Vecteurs aléatoires	184
D.1.4.1	Lois de probabilité	184
D.1.4.2	Fonctions caractéristiques, moments et cumulants	185
D.1.4.3	Matrice de covariance et de corrélation	185
D.1.5	Estimations	186
D.1.5.1	Echantillon, statistique et estimateur	186
D.1.5.2	Estimation des moments et des cumulants	186
D.2	Application aux signaux temporels	187
D.2.1	Caractérisation des processus aléatoires	187
D.2.1.1	Définitions préliminaires	187
D.2.1.2	Propriétés statistiques des deux premiers moments	188
D.2.2	Ergodicité et stationnarité	189
D.2.2.1	Ergodicité	189
D.2.2.2	Stationnarité	189
D.2.2.3	Ergodicité et stationnarité	190
D.2.3	Moments temporels d'ordres un et deux d'un signal stationnaire	190
D.2.4	Multicorrélations	192
D.2.4.1	Temps continu	192
D.2.4.2	Temps discret	193
D.2.4.3	Estimation	193
E	Choix du paramètre de dispersion en classification spectrale	195
F	Simulations sur une base réduite de thrombi	197
F.1	Protocole expérimental	198
F.2	Echogénéicité du thrombus	199
F.2.1	Echographie transversale	199
F.2.2	Echographie longitudinale	201
F.3	Dureté du thrombus	202
F.3.1	Elastographie transversale	202
F.3.2	Elastographie longitudinale	203
F.4	Echogénéicité et dureté du thrombus	204
F.4.1	Echographie et élastographie transversales	204
F.4.2	Echographie et élastographie longitudinales	205
F.5	Bilan des simulations sur les images de thrombi	205
G	Simulations complémentaires sur la base de thrombi	207
G.1	Détection d'un cancer	207
G.2	Incidence du tabagisme	209
G.3	Impact des paramètres d'acquisition des échographies	210
G.3.1	Modes fondamental et harmonique	210
G.3.2	Modes d'imagerie et gain	211
G.3.3	Conclusion sur l'influence des paramètres d'acquisition	213
	Bibliographie	215

Table des figures

1.1	Système veineux des membres inférieurs.	5
1.2	De la thrombose veineuse à l'embolie pulmonaire.	7
1.3	Diagnostic par échographie de la thrombose veineuse profonde.	8
1.4	Ondes acoustiques de compression et de cisaillement.	9
1.5	Distorsion d'une onde acoustique dans un milieu non linéaire.	11
1.6	Echographie en mode A ou B, unidimensionnelle ou bidimensionnelle.	11
1.7	Principe des trois modes d'échographie : fondamental, THI et DTHI.	12
1.8	Comparatif des trois modes d'échographie : fondamental, THI et DTHI.	13
1.9	Déformation élastique des tissus biologiques sous l'action d'une contrainte externe. . .	14
1.10	Principe de l'élastographie dynamique par force de radiation.	15
1.11	Echographie et Elastographie avec le système Toshiba Aplio 500.	16
1.12	Description d'une acquisition obtenue avec le système Toshiba Aplio 500	17
1.13	Principaux types de sondes : linéaires, convexes, et sectorielles.	18
1.14	Protocole d'acquisition des échographies et des élastographies	22
1.15	Echographies et élastographies en mode transversal et longitudinal.	22
1.16	Echographies et élastographies de glandes salivaires.	26
1.17	Base de glandes salivaires.	27
2.1	Techniques d'élastographie dynamique existantes.	32
2.2	Chaîne de traitement d'images.	34
2.3	Illustration du <i>Manifold untangling</i>	37
2.4	Rétrécissement et agrandissement d'une image par TFD et TCD.	41
2.5	Projection de glandes salivaires sur deux axes par ACP.	41
2.6	Réduction la dimension en préservant les distances.	42
2.7	Principe de la classification supervisée.	44
2.8	Astuce du noyau d'un SVM.	45
2.9	Principe d'un SVM.	46
2.10	Réseau de neurones organisé en quatre couches.	47
2.11	Architecture d'un réseau de neurones.	47
2.12	Itérations successives de l'algorithme des K -moyennes en deux classes.	49
2.13	Itérations successives de l'algorithme DBSCAN en deux classes.	50
2.14	Exemple de graphe pondéré.	51
2.15	Droite et plan de projection.	54
3.1	Extraction des différentes images pour une acquisition.	59
3.2	Extraction des différentes images pour une acquisition.	60
3.3	IHM <i>Lectio</i> pour lire et convertir les fichiers DICOM	62
3.4	Rognage des images à partir des métadonnées.	63
3.5	Echographies et histogrammes avant et après égalisation.	64
3.6	Conversion de l'élastogrphie couleur en carte de vitesses.	65
3.7	IHM <i>Moasic</i> pour segmenter les échographies de thrombus.	67
3.8	Exemple d'imagettes à partir des échographies de thrombus.	68
3.9	Segmentation et extraction d'imagettes sur les échographies de glandes salivaires. . . .	68

3.10 IHM <i>Optio</i> pour extraire manuellement des imagettes.	69
4.1 Contraction de l'espace par l'opérateur Φ pour obtenir une meilleure représentation	72
4.2 Support fréquentielle d'une ondelette mère et des ondelettes filles.	73
4.3 Représentation du SO en réseau de convolutions.	75
4.4 Représentation des coefficients du scattering dans le plan fréquentiel.	76
4.5 Représentation des scattering coefficients d'une image composée de rayures.	76
4.6 Autre représentation des scattering coefficients d'un image composée de rayures.	77
4.7 Intérêt de l'invariance en rotation.	78
4.8 Autocorrélation d'une image.	81
4.9 Corrélacion croisée de deux images.	82
4.10 Bicorrélation d'une image.	82
4.11 Tricorrélation d'une image.	83
4.12 Axes choisis pour les multicorrélations.	85
4.13 Structure des descripteurs du SO et des SOS.	86
5.1 Classes de textures de la base KTH-TIPS.	91
5.2 Procédure de validation croisée en 9 sous-ensembles.	92
5.3 Résultats de la procédure de validation croisée.	93
5.4 Meilleurs paramètres du scattering operator.	93
5.5 Logarithme, invariances et apprentissage multi-échelles.	94
5.6 Evaluation du scattering operator sur la base KTH-TIPS.	96
5.7 <i>Manifold untangling</i> du <i>scattering operator</i>	98
5.8 Pouvoir discriminant des multicorrélations en fonction de leur ordre.	100
5.9 <i>Manifold untangling</i> des multicorrélations en fonction de l'ordre.	101
5.10 Multicorrélations d'ordre 2 en fonction des axes sélectionnés.	102
5.11 Multicorrélations d'ordre supérieur en fonction des axes et des ordres sélectionnés.	102
5.12 <i>Manifold untangling</i> des multicorrélations en fonction des axes choisis.	103
5.13 Uniformisation des descripteurs multirésolutions et statistiques.	105
5.14 Classification des descripteurs multirésolutions et statistiques uniformisés.	106
5.15 <i>Manifold untangling</i> des descripteurs multirésolutions et statistiques.	106
5.16 Imagettes d'une glande parotide tirées d'une échographie et d'une élastographie.	109
5.17 Descripteurs multirésolutions et statistiques des imagettes.	110
5.18 Protocole expérimental des simulations sur les glandes salivaires.	112
5.19 <i>Manifold untangling</i> de la base de glandes salivaires.	118
5.20 <i>Manifold untangling</i> des élastographies des glandes parotides.	121
5.21 Base de thrombi organisée en quatre classes.	125
A.1 Mécanisme assurant l'arrêt du saignement en cas de liaison d'un vaisseau.	149
A.2 Clou plaquettaire : adhésion plaquettaire et agrégation plaquettaire.	150
A.3 Echogénéicité d'un thrombus et détection d'un thrombus par doppler veineux.	151
B.1 Evolution de l'IHM de lecture et conversion des fichiers DICOM	154
B.2 Suppression des annotations et de l'ellipse.	155
B.3 Segmentation des ellipses.	156
B.4 Cas 3 : agrandissement de l'ellipse.	157
B.5 Cas 4 : dessin de l'ellipse.	157
B.6 Cas 4 : dessin de l'ellipse.	158
B.7 Extraction d'une image carrée la plus grande possible à l'intérieur d'une ellipse	158
C.1 Analyse d'un signal composée de trois sinusoides.	162
C.2 Représentation des ondelettes de Haar, du chapeau mexicain et de Morlet.	164
C.3 Décomposition de Haar d'une fonction en escalier.	166
C.4 Structure de décomposition en ondelettes sur 3 niveaux de résolution.	168
C.5 Analyse multirésolution M -bandes sur 2 niveaux de résolution.	169

C.6	Structure de décomposition en paquets d'ondelettes sur 3 niveaux de résolution.	169
C.7	Structure d'un banc de filtres d'analyse en ondelettes séparables 2D.	171
C.8	Représentation d'une AMR en ondelettes séparables sur 3 niveaux de résolutions.	172
C.9	AMR en ondelettes séparables sur 3 niveaux de résolutions d'une image.	172
C.10	AMR en ondelettes 3-bandes et en paquet d'ondelettes.	173
C.11	Transformée en ondelettes 2D directionnelles et omnidirectionnelles.	174
C.12	Balayage fréquentiel par des ondelettes 2D.	174
D.1	Densité de probabilité d'une variable gaussienne centrée et réduite	178
D.2	Intérêt de l'invariance en rotation.	182
D.3	Principe de l'autocorrélation.	191
D.4	Bicorrélation et tricorrélation du chapeau mexicain.	193
E.1	Optimisation du paramètre de dispersion en classification spectrale.	195

Liste des tableaux

1.1	Caractéristiques acoustiques de certains milieux.	10
1.2	Liste des patients avec le protocole 1.	21
1.3	Liste des acquisitions par patient obtenu avec le protocole 2.	23
1.4	Paramètres d’acquisition du protocole 2	24
1.5	Liste des acquisitions par patient obtenu avec le protocole 3.	25
2.1	Approches de débruitage d’images échographiques (« speckle reduction »).	35
2.2	Approches de segmentation d’images.	36
2.3	Approches d’extraction de descripteurs.	37
2.4	Approches de classification des données.	38
2.5	Calcul du Laplacien non normalisé.	52
5.1	Matrices de confusion obtenues avec le SO et la SMA.	95
5.2	Comparaison des performances du SO et de trois autres approches.	95
5.3	Matrices de confusion obtenues avec les k -moyennes et la classification spectrale.	97
5.4	Méthodes d’uniformisation des descripteurs multirésolutions et statistiques.	104
5.5	Echelle des couleurs en fonction des pourcentages de classification.	111
5.6	Classification par SMA avec 50 % des glandes consacrées à l’apprentissage.	114
5.7	Classification des glandes par SMA en diminuant la proportion d’apprentissage.	115
5.8	Classification pour deux algorithmes de partitionnement.	116
5.9	Classification spectrale des descripteurs de glandes.	117
5.10	Matrice de confusion, sensibilité et spécificité pour les glandes salivaires.	118
5.11	Classification des glandes parotides.	120
5.12	Classification des glandes submandibulaires.	122
5.13	Matrice de confusion par SMA en fonction des facteurs déclencheurs.	126
5.14	Matrice de confusion par k -moyennes en fonction des facteurs déclencheurs.	127
5.15	Matrice de confusion par CS en fonction des facteurs déclencheurs.	127
5.16	Classification des thrombi veineux en fonction des facteurs déclencheurs.	128
5.17	Matrice de confusion par SMA : TVP provoquée ou idiopathique, avec ou sans EP.	129
5.18	Classification vis-à-vis du caractère provoqué de la TVP et de la présence d’EP.	130
5.19	Matrice de confusion par k -moyennes : TVP provoquée/idiopathique, avec/sans EP.	131
5.20	Matrice de confusion par CS : TVP provoquée ou idiopathique, avec ou sans EP.	131
5.21	Matrice de confusion par SMA entre TVP provoquée et idiopathique.	132
5.22	Matrice de confusion par k -moyennes entre TVP provoquée et idiopathique.	132
5.23	Matrice de confusion par CS entre TVP provoquée et idiopathique.	132
5.24	Classification des thrombi par rapport au caractère provoqué de la TVP.	133
5.25	Matrice de confusion par SMA en fonction de la présence d’EP.	134
5.26	Matrice de confusion par k -moyennes en fonction de la présence d’EP.	134
5.27	Matrice de confusion par CS en fonction de la présence d’EP.	134
5.28	Classification des thrombi suivant la présence d’EP.	135
F.1	Liste des acquisitions par patient obtenues avec le protocole 2B	198
F.2	Classification des échographies transversales.	199
F.3	Classification des échographies longitudinales.	201

F.4	Classification des élastographies transversales.	202
F.5	Classification des élastographies longitudinales.	203
F.6	Classification des échographies/élastographies transversales.	204
F.7	Classification des échographies/élastographies en mode longitudinal.	205
F.8	Similitudes entre la classification des images en mode transversal et longitudinal.	206
F.9	Similitudes entre la classification des échographies et élastographies.	206
G.1	Classification des thrombi suivant leur association à un cancer.	208
G.2	Matrice de confusion par CS en fonction de la présence d'un cancer.	208
G.3	Matrice de confusion par SMA en fonction de la présence d'un cancer.	208
G.4	Classification des thrombi par rapport au tabagisme.	209
G.5	Matrice de confusion par CS par rapport au tabagisme.	209
G.6	Matrice de confusion par SMA par rapport au tabagisme.	209
G.7	Classification des thrombi par rapport au mode d'imagerie échographique.	210
G.8	Matrice de confusion par SMA par rapport au mode d'imagerie échographique.	211
G.9	Matrice de confusion par CS par rapport au mode d'imagerie échographique.	211
G.10	Matrice de confusion par SMA par rapport au mode et au gain.	211
G.11	Matrice de confusion par CS par rapport au mode et au gain.	212
G.12	Classification des thrombi par rapport au mode d'imagerie échographique et au gain.	212

Glossaire médical

Terme	Description	Pages
E	Module de Young qui caractérise la dureté d'un solide (en Pa)	3, 13
K	Modules élastiques de compression (en Pa)	3, 9
ϵ	Déformation d'un solide par une contrainte externe statique (adimensionnel)	3, 13, 14
μ	Modules élastiques de cisaillement (en Pa)	3, 9
ρ	Densité volumique de masse (ou masse volumique) d'un milieu (en kg/m^3)	3, 9, 10, 14
σ	Contrainte externe statique appliquée sur la surface d'un solide (en Pa)	3, 13
2D	Deux Dimensions	29, 31, 46, 70, 79, 88, 97, 99, 102, 106, 169, 170, 173
3D	Trois Dimensions	29, 31
A	Indice de qualité « Aplipure » : paramètre du système Toshiba Aplio 500 pour ajuster la qualité des échographies	3, 20, 24
BDD	Base De Données	152
DICOM	Format standard pour la gestion des images biomédicales (<i>Digital Imaging and Communications in Medicine</i>)	3, 17, 19, 56, 60, 61, 64, 140, 152, 153
DTHI	Differential Tissue Harmonic Imaging : échographie des tissus humains basée sur l'analyse de la différence des harmoniques de deux fréquences	3, 12, 13, 16, 19, 21, 23, 56, 63
EP	Embolie Pulmonaire	3, 6, 7, 21, 24, 29, 55, 56, 61, 88, 89, 123, 124, 128–130, 133, 134, 144, 196, 198, 199, 203, 204, 206
FGS	<i>Fondamental Gray Scale</i> : échographie formée à partir de la fondamentale des ultrasons réfléchis	3, 24, 206, 209, 210
FN	Faux Négatifs c'est-à-dire le nombre (ou le pourcentage) de patients malades classés comme normals.	88, 107
FP	Faux Positifs (ou fausse alarme) c'est-à-dire le nombre (ou le pourcentage) de patients sains classés comme malades.	88, 107

Terme	Description	Pages
ID	Identifiant	56, 62, 88, 123, 124
IRM	Imagerie par Résonance Magnétique	3, 8, 29, 30, 33, 34, 51
Mode A	Echographie en mode Amplitude	3, 11, 12
Mode B	Echographie en mode Brillance	3, 11, 19, 20, 29, 32
MTVE	Maladie Veineuse ThromboEmbolique	3, 6
P	Indice de qualité « Precision + » : paramètre du système Toshiba Aplio 500 pour ajuster la qualité des échographies	3, 20, 24
RGB	Codage des couleurs en rouge, vert et bleu (<i>Red, Green, Blue</i> en anglais)	56, 57, 64
ROI	Région d'intérêt (<i>Region Of Interest</i>)	3, 17, 19, 29, 32, 56, 65, 70
SWIRE	Shear Wave Induced Resonance Elastography	29, 33
TEP	Tomographie par Emission de Positons	29, 34
THI	Tissue Harmonic Imaging : échographie des tissus humains formée à partir des harmoniques des ultrasons réfléchis	3, 12, 13, 16, 19, 21, 24, 56, 57, 63, 206, 209, 210
TVP	Thrombose Veineuse Profonde	3, 4, 7–9, 21, 23, 25, 29, 30, 32, 55, 88, 123–126, 128–133, 137, 196, 198, 199, 204, 206, 208
VN	Vrais Négatifs c'est-à-dire le nombre (ou le pourcentage) de patients sains classés comme normals.	88, 107
VP	Vrais Positifs c'est-à-dire le nombre (ou le pourcentage) de patients malades classés correctement.	88, 107

Glossaire du traitement d'images

Terme	Description	Pages
ACP	Analyse en Composantes Principales (ou <i>Principal Component Analysis</i> en anglais, PCA)	xii, 29, 41–44, 46, 50, 55, 70, 82, 88, 91, 144, 177, 194, 196, 197, 206–212
AME	Apprentissage Multi-Echelle (ou multirésolution)	70, 78, 88, 90, 93, 94
AMR	Analyse MultiRésolution (ou multi-échelle)	70, 71, 73, 164, 166, 169–172
CLAHE	Egalisation d'histogramme locale : <i>Contrast Limited Adaptive Histogram Equalization</i>	56, 63, 88, 109
CS	Classification Spectrale (<i>spectral clustering</i> en anglais)	29, 51, 53, 88, 89, 95–97, 110, 112, 115–118, 120, 122, 131, 196, 197, 199, 200, 206, 208, 209
DBSCAN	Density-Based Spatial Clustering of Applications with Noise est un algorithme de partitionnement spatial à base de densité.	29, 38, 44, 50, 51, 55, 88, 89, 97–100, 110–112, 114, 115, 125, 141
EH	Egalisation d'Histogramme	29, 35, 36, 55
IHM	Interface Homme-Machine	56, 60, 61, 64, 66–69, 140, 143, 152–154, 156
KTH-TIPS	Base de textures créée par des membres de l'Institut royal de technologie de Stockholm (<i>Kungliga Tekniska Högskolan</i>) et contenant dix classes de matériaux observés sous différentes luminosités, orientations et échelles (<i>Textured under varying Illumination, Pose and Scale</i>) [56].	x, 88–90, 92–94, 96–99, 102–104, 109–111, 114, 118, 144, 145, 194, 196, 197

Terme	Description	Pages
MAP	Estimateur statistique du Maximum <i>A Posteriori</i> : méthode pour estimer des paramètres inconnus d'un modèle à partir de la vraisemblance et de la distribution <i>a priori</i> des paramètres.	29, 35
MV	Estimateur statistique du Maximum de Vraisemblance : méthode pour estimer des paramètres inconnus d'un modèle en construisant une fonction de vraisemblance (déduite de la fonction de densité) puis en maximisant son logarithme par rapport aux paramètres inconnus. Le MV est l'estimateur MAP pour une distribution <i>a priori</i> uniforme.	29, 32
NGM	Niveau de Gris Médian	29, 30
NJW	Algorithme de Ng, Jordan et Weiss	29, 54
OPTICS	Ordering Points to Identify the Clustering Structure est un algorithme de partitionnement spatial à base de densité.	29, 38, 88, 114
RN	Réseau de Neurones	29, 46–48
ROC	Courbe permettant de mesurer la performance d'un classifieur (<i>Receiver Operating Characteristic</i> en anglais)	29, 32
SMA	Classification par Sélection de Modèles Affines	29, 38, 44, 45, 55, 88, 89, 91, 93, 94, 96, 97, 99, 100, 104, 110, 112–114, 116, 117, 119, 121, 125–135
SO	<i>Scattering Operator</i> : réseau convolutif de transformée en ondelettes et d'opérateurs contractants	v, xii, 29, 55, 70–78, 85, 86, 88–96, 103–106, 109, 111, 112, 114, 115, 117, 118, 133, 136, 194, 196, 197, 199–201, 203, 205, 211
SOS	Statistiques d'Ordre Supérieur (<i>High Order Statistics</i> en anglais, HOS)	v, 70, 85, 86, 88, 89, 103–106, 109, 111, 112, 115, 136
SVM	Machine à vecteurs de support ou séparateur à vaste marge (<i>Support Vector Machine</i> en anglais)	ix, 29, 38, 44–46, 55
t-SNE	t-distributed Stochastic Neighbor Embedding	29, 42, 43, 55, 88, 97, 99, 100, 102, 116, 117, 120
TCD	Transformée en Cosinus Discrète	29, 37, 40, 41, 43, 152, 158, 196, 197
TFD	Transformée de Fourier Discrète	29, 37, 40, 41, 70, 81, 152, 158

Introduction

Introduction générale

La thèse s'est déroulée à l'ENSTA Bretagne en partenariat avec l'Unité d'Echo-doppler et de Médecine Vasculaire du CHU de Brest. En principe, un thrombus veineux se forme dans un vaisseau sanguin blessé pour limiter le saignement le temps de sa guérison. De nombreux facteurs isolés ou associés, acquis ou génétiques, peuvent ralentir le flux sanguin, altérer les parois veineuses et/ou troubler la coagulation-fibrinolyse et ainsi générer un thrombus. Ce phénomène apparaît le plus souvent dans le réseau veineux profond des membres inférieurs. Dans ce cas, on parle de thrombose veineuse profonde ou, lorsqu'il y a une inflammation de la veine autour du thrombus, de phlébite (du grec *phlebos* qui signifie la veine et du suffixe médical *ite* qui indique l'inflammation [17]). En plus des problèmes liés à la mauvaise circulation du sang, un fragment ou la totalité du thrombus peut se détacher, migrer jusqu'au cœur et obstruer l'artère pulmonaire. Cette complication majeure, pouvant entraîner une mort rapide, est nommée embolie pulmonaire et touche environ 430 mille personnes chaque année en Europe [36].

L'enjeu de notre projet est de caractériser la structure du thrombus veineux afin d'identifier un ou plusieurs facteurs responsables de sa formation. La possibilité de prédire le risque d'embolie pulmonaire serait également une avancée majeure. Comparés aux thrombi artériels (accident vasculaire cérébral et infarctus du myocarde notamment), les thrombi veineux sont moins bien connus de la population [32] et sont également moins étudiés par la communauté scientifique, ce qui confère à notre projet toute son originalité. Actuellement, deux types d'imagerie sont majoritairement considérés pour observer un thrombus veineux : l'échographie et l'élastographie. Ces dernières utilisent les propriétés de propagation des ultrasons dans les tissus humains pour cartographier respectivement leur échogénicité, et leur dureté. A partir des images obtenues suivant ces deux modes d'imagerie, nous proposons de les prétraiter, de sélectionner des régions d'intérêt puis d'extraire des attributs de ces régions (appelés descripteurs) afin de caractériser la structure des thrombi. Ensuite, les attributs extraits des images sont analysés en fonction de l'origine de la thrombose et de la présence (ou non) d'embolie pulmonaire.

Organisation du manuscrit

Au cours de cette thèse, nous avons adapté un ensemble d'outils du traitement du signal et du traitement d'images à des données médicales : des échographies et des élastographies de thrombi veineux. Pour caractériser leur structure, nous avons mis en place une chaîne de traitement d'images complète, du prétraitement des données jusqu'à l'étape de classification. L'enjeu du projet est, d'à travers cette chaîne, d'identifier si la structure du thrombus est différente suivant les facteurs responsables de sa formation et suivant sa propriété à se disloquer et à migrer (ou non) vers les artères pulmonaires (*i.e.* embolie pulmonaire). Pour répondre à cette problématique, nous proposons d'analyser les performances de classification des images échographiques et élastographiques pour différents paramétrages de la chaîne de traitement. Ce manuscrit résume donc les travaux réalisés durant les trois années de cette thèse et s'articule de la façon suivante :

- Le chapitre 1 présente le contexte médical et technique de la thèse et l'évolution du protocole de collecte des données par l'Unité d'Echo-doppler et de Médecine Vasculaire du CHU de Brest. Ce chapitre permet de mieux comprendre les enjeux de la thèse et ses complexités.

- Le chapitre 2 est consacré à la problématique de caractérisation et la classification de régions d'intérêt en imagerie biomédicale. Il abordera tout d'abord les premières études couplant échographie et élastographie en imagerie du thrombus puis présentera de manière plus générale les principales approches employées en traitement du signal.

- Le chapitre 3 présente les bases de données utilisées pour nos simulations et l'ensemble des étapes de prétraitement préalable à la caractérisation et à la classification. Pour leur validation, les approches considérées ont d'abord été testées sur une base d'images de textures puis sur une base d'images de glandes salivaires. Ce chapitre décrira également la préparation des échographies et des élastographies ainsi que la sélection des régions d'intérêt.

- Le chapitre 4 décrit les deux approches développées pour caractériser et classifier nos images de thrombi : le scattering operator et les statistiques d'ordre supérieur. Le scattering operator s'appuie sur la théorie des ondelettes pour rechercher des éléments caractéristiques à différents « zooms » de l'image. Comme le ferait l'œil humain pour observer un tableau, un paysage ou une photo, cette méthode commence par s'intéresser à l'image dans sa globalité puis à ses différents niveaux de détails (formes, objets, contours, *etc.*). Les statistiques d'ordre supérieur se basent sur la théorie des probabilités. La plupart des méthodes usuelles sont fondées sur une description simplifiée d'un signal aléatoire (ou d'une variable aléatoire) faisant appel aux moments d'ordres un (moyenne) et deux (écart-type, variance, corrélation). Dans ce chapitre, nous nous intéresserons également aux ordres supérieurs à deux qui, par leur plus grand degré de généralité, permettent d'obtenir des informations complémentaire qui peuvent être utilisées pour renforcer la classification des données.

- Enfin, le chapitre 5 est consacré aux simulations et à leurs résultats. Nous évaluerons tout d'abord les approches décrites au chapitre précédent sur les bases de tests (textures et glandes salivaires) puis elles seront appliquées sur nos images de thrombi. Nous observerons alors si leur classement peut être lié aux facteurs responsables de leur formation et/ou à la survenue (ou l'absence) d'une embolie pulmonaire.

- Conclusion générale et perspectives

Chapitre 1

Contexte et problématiques

Glossaire du chapitre 1

E Module de Young qui caractérise la dureté d'un solide (en Pa)

K Modules élastiques de compression (en Pa)

ϵ Déformation d'un solide par une contrainte externe statique (adimensionnel)

μ Modules élastiques de cisaillement (en Pa)

ρ Densité volumique de masse (ou masse volumique) d'un milieu (en kg/m^3)

σ Contrainte externe statique appliquée sur la surface d'un solide (en Pa)

A Indice de qualité « Aplipure » : paramètre du système Toshiba Aplio 500 pour ajuster la qualité des échographies

DICOM Format standard pour la gestion des images biomédicales (*Digital Imaging and Communications in Medicine*)

DTHI Differential Tissue Harmonic Imaging : échographie des tissus humains basée sur l'analyse de la différence des harmoniques de deux fréquences

EP Embolie Pulmonaire

FGS *Fondamental Gray Scale* : échographie formée à partir de la fondamentale des ultrasons réfléchis

IRM Imagerie par Résonance Magnétique

MTVE Maladie Veineuse ThromboEmbolique

Mode A Echographie en mode Amplitude

Mode B Echographie en mode Brillance

P Indice de qualité « Precision + » : paramètre du système Toshiba Aplio 500 pour ajuster la qualité des échographies

ROI Région d'intérêt (*Region Of Interest*)

THI Tissue Harmonic Imaging : échographie des tissus humains formée à partir des harmoniques des ultrasons réfléchis

TVP Thrombose Veineuse Profonde

Introduction

Dans ce chapitre, nous revenons plus en détail sur le contexte de la thèse, tant d'un point de vue médical que technique et matériel. La section 1.1 s'intéresse à la formation du thrombus veineux et aux possibles conséquences pour le patient. Ensuite, la section 1.2 introduit les différents moyens techniques à notre disposition pour analyser la structure du thrombus. Enfin, la section 1.3 présente la base de données que nous avons construite depuis deux ans.

1.1 Thrombose veineuse profonde

Cette section décrit l'aspect médical de la thèse et notamment les phénomènes responsables de la formation, adaptée ou non, d'un thrombus veineux. Après avoir considéré les principales causes de thrombose, la suite de cette section s'intéresse aux principaux risques encourus par le patient, au diagnostic et aux traitements de la pathologie.

1.1.1 Formation adaptée d'un thrombus veineux

En temps normal, un thrombus se développe en cas de blessure d'un vaisseau sanguin pour limiter le saignement [77]. Ce phénomène, qui transforme le sang de l'état liquide à l'état plus ou moins solide, est appelé la coagulation sanguine. Un thrombus est constitué principalement de plaquettes (petits éléments du sang) et de fibrinogènes (protéine soluble synthétisée par le foie). Au cours de la coagulation, l'enchaînement de réactions chimiques, impliquant divers substrats et enzymes plasmatiques, transforme le fibrinogène en une protéine insoluble, la fibrine, qui forme l'armature du thrombus (voir annexe A.1 pour davantage de précisions). Le thrombus devient donc de plus en plus solide au cours du temps. Une fois le vaisseau réparé, le thrombus est dissout afin de rétablir la circulation sanguine. Ce processus indispensable, nommé fibrinolyse, élimine les déchets de fibrine afin de rendre le vaisseau de nouveau perméable (c'est-à-dire que le sang peut à nouveau circuler normalement dans les vaisseaux libres de tout thrombus).

1.1.2 Formation inadaptée

De nombreux mécanismes physiopathologiques peuvent venir déséquilibrer la circulation sanguine et engendrer la formation ou la non-dissolution d'un thrombus. Une Thrombose Veineuse Profonde (TVP) correspond à l'occlusion partielle ou totale d'une veine profonde (par opposition avec une veine en surface dite superficielle). Elle commence le plus souvent dans les veines du mollet et s'étend progressivement vers les veines de la cuisse. Une thrombose veineuse, associée à une inflammation des veines, est plus communément appelée phlébite. Selon la triade décrite par Virchow en 1956 [5], trois types de phénomènes, associés ou non, favorisant la survenue de thrombose : la stase sanguine, les lésions des veines et la thrombophilie.

1.1.2.1 La stase sanguine

Le sang circule à sens unique dans un système clos formé de différents vaisseaux (*cf.* figure 1.2.a). Les artères transportent le sang oxygéné du cœur vers les divers organes grâce aux contractions rythmiques du muscle cardiaque. Les veines prennent ensuite le relais pour transporter le sang appauvri en oxygène des organes vers le cœur. Le battement cardiaque, la respiration abdominale, la contraction des muscles du mollet et l'écrasement de la voûte plantaire permettent de faire remonter le sang des pieds vers les veines [77]. La stase veineuse, c'est-à-dire le ralentissement du flux veineux, favorise la formation d'un thrombus, notamment au niveau des membres inférieurs dont le réseau veineux est illustré en figure 1.1.a.

Une immobilisation prolongée due à une maladie, un alitement ou un accouchement limite l'action des mollets et de la voûte plantaire sur la circulation (voir figure 1.1.b) et peut donc favoriser l'apparition de thrombose. Pour prévenir les risques de thrombose, les immobilisations plâtrées des membres inférieurs peuvent par exemple être accompagnées d'une injection quotidienne d'un anticoagulant par voie sous-cutanée. En outre, une station assise prolongée, un voyage long en avion ou un cancer peuvent également causer un ralentissement du flux sanguin. Une étude [95] a montré que le risque de thrombose lors d'un vol d'une durée supérieure à 5 heures est estimé entre 2 et 4 pour 10 000 passagers. De la même manière, l'insuffisance cardiaque et les varices des membres inférieurs sont des facteurs possibles pour générer des stases veineuses.

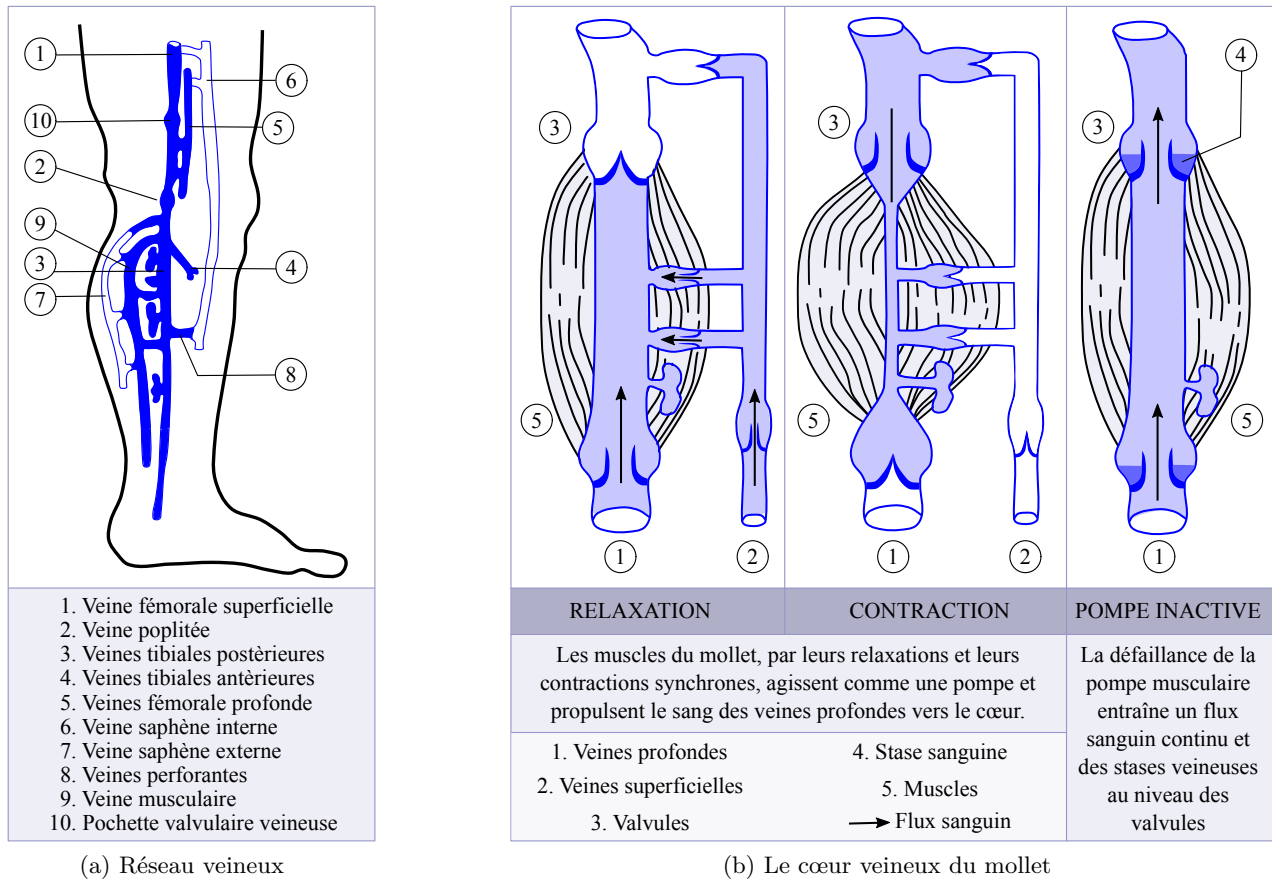


FIGURE 1.1 – Système veineux des membres inférieurs (inspiré de [77]).

1.1.2.2 L'altération des parois veineuses

Les veines sont dotées de systèmes de valvules pour éviter le reflux sanguin et possèdent des parois beaucoup plus fines et élastiques que les artères. Elles sont donc plus sensibles à la chirurgie, aux traumatismes et au vieillissement qui peuvent les détériorer et donc limiter leur efficacité. De plus, en cas d'infection, il se produit une inflammation des veines et donc des retours veineux pouvant causer des phlébites. La survenue d'une thrombose veineuse après la sortie de l'hôpital est d'autant plus probable que la chirurgie et l'âge du patient sont importants. Les études phlébographiques [95] montrent une prévalence de 10 à 20 % de nouvelles thromboses veineuses dans les 4 à 5 semaines après la sortie de l'hôpital sans prophylaxie (i.e. sans mesure visant à empêcher l'apparition d'une phlébite). L'incidence des thromboses veineuses augmente de façon exponentielle avec l'âge : le risque double tous les dix ans à partir de 40 ans [16, 95, 144].

1.1.2.3 La thrombophilie

La thrombophilie regroupe l'ensemble des mécanismes, acquis ou génétiques, responsables d'un renforcement de la coagulation physique ou d'un déficit de l'anticoagulation. Les protéines C et S, synthétisées par le foie, agissent conjointement pour inhiber la coagulation du sang. Des déficits en protéines C ou S, ou des dysfonctionnements d'une des deux protéines peuvent conduire à une activation excessive de la coagulation et donc provoquer des thromboses veineuses. Par exemple, la grossesse (au troisième trimestre) peut réduire les taux de ces deux protéines et, à l'inverse, un traitement œstrogénique peut entraîner une augmentation de ces taux : le risque de thrombose deviendrait de 2 à 5 fois plus important [95] par rapport à une personne non enceinte ou sous aucun traitement hormonal.

Les anomalies génétiques (facteur V Leiden, gène FII G20210A) ainsi que certaines situations acquises (obésité, maladies du foie, du rein, infections graves, cancers [57], *etc.*) peuvent entraîner une diminution de la production ou une augmentation de la consommation de ces protéines et donc favoriser l'apparition de thromboses [77]. Enfin, des études encore controversées montreraient une incidence de la cigarette sur les thromboses veineuses [15, 26, 64]. En effet, le tabagisme, en augmentant le taux de fibrinogène dans le sang, favorise l'agrégation des plaquettes et donc la coagulation. Parallèlement, pour compenser le manque d'oxygène, les globules rouges et blancs réagissent en augmentant leur nombre rendant ainsi le sang plus visqueux. Ces deux phénomènes couplés augmenteraient ainsi le risque de thrombose chez le fumeur.

1.1.3 Observation clinique et risques évolutifs

Cette section s'appuie sur le Tome 2 du *Traité de médecine vasculaire* sur les maladies veineuses, lymphatiques et microcirculatoires [49] et sur le mémoire de DESC (Diplôme d'Etude Spécialisé Complémentaire) de médecine vasculaire de S. Beuil [14].

Un thrombus met environ 10 à 15 jours pour se former et le patient consulte un médecin au plus tôt aux premiers signes évidents de phlébite. D'après une étude Ifop/Bayer de 2014 sur un échantillon de 1005 personnes représentatif de la population française âgée de 18 ans ou plus [32], 60 % des personnes interrogées ne connaissent pas les trois principaux signes : douleur sourde au mollet ou à la cuisse, sensation de chaleur et gonflement (œdème) du mollet. Les symptômes dépendent de l'importance de la réaction inflammatoire autour du thrombus et de sa taille. Lorsqu'une veine profonde est obstruée par le thrombus et que le sang ne peut plus remonter par cette veine, il va se former une hyperpression en amont du thrombus et le flux sanguin va alors emprunter le réseau superficiel. Il n'y a parfois aucun symptôme rendant ainsi la maladie d'autant plus dangereuse.

Outre les trois principaux signes exposés ci-dessus, il est possible d'observer, au stade initial, d'autres symptômes locaux comme des engourdissements, des crampes, des jambes lourdes. On peut également observer des signes cliniques généraux comme un pouls rapide avec une température normale, des angoisses, des sensations de malaise inexplicables, une douleur à la palpation le long du trajet veineux, le signe Homans (douleur à la dorsiflexion du pied), une dilatation asymétrique des veines superficielles et une peau brillante et dure, blanche ou bleuâtre. Au stade de la thrombose confirmée, les douleurs risquent de s'intensifier, pouvant aller jusqu'à l'impotence fonctionnelle complète, et s'accompagner de fièvre, de lésions cutanées, de tachycardie et d'œdème cylindrique (jambe en poteau).

Les risques évolutifs majeurs sont la maladie post-phlébitique qui peut altérer significativement la qualité de vie et l'embolie pulmonaire (voir figure 1.2) qui met en jeu le pronostic vital. La première complication est chronique et associée à des degrés divers une sensation de lourdeur des jambes, des varices, des œdèmes, des troubles trophiques pouvant aller jusqu'à un ulcère de jambe. La seconde complication se produit en cas de décollement du thrombus et de sa migration vers le cœur. On parle d'Embolie Pulmonaire (EP) ou Maladie Veineuse ThromboEmbolique (MTVE) lorsque le thrombus se bloque, plus ou moins tôt selon sa taille, dans une artère pulmonaire. Un tel événement peut ne manifester aucun symptôme mais peut aussi se traduire comme un point de côté, une dyspnée, une détresse respiratoire, des crachats de sang ou, dans le pire des cas, une mort subite.

La thrombose veineuse profonde ne s'accompagne pas systématiquement d'une embolie pulmonaire. D'une manière générale, un patient faisant une phlébite non accompagnée d'une embolie pulmonaire a une forte probabilité de récidiver. Ensuite, un patient qui fait une embolie pulmonaire lors de son premier épisode de phlébite va, s'il récidive, le plus souvent refaire une embolie pulmonaire. Le risque de récidive après une thrombose veineuse est évalué entre 5 et 10 % par an [95].

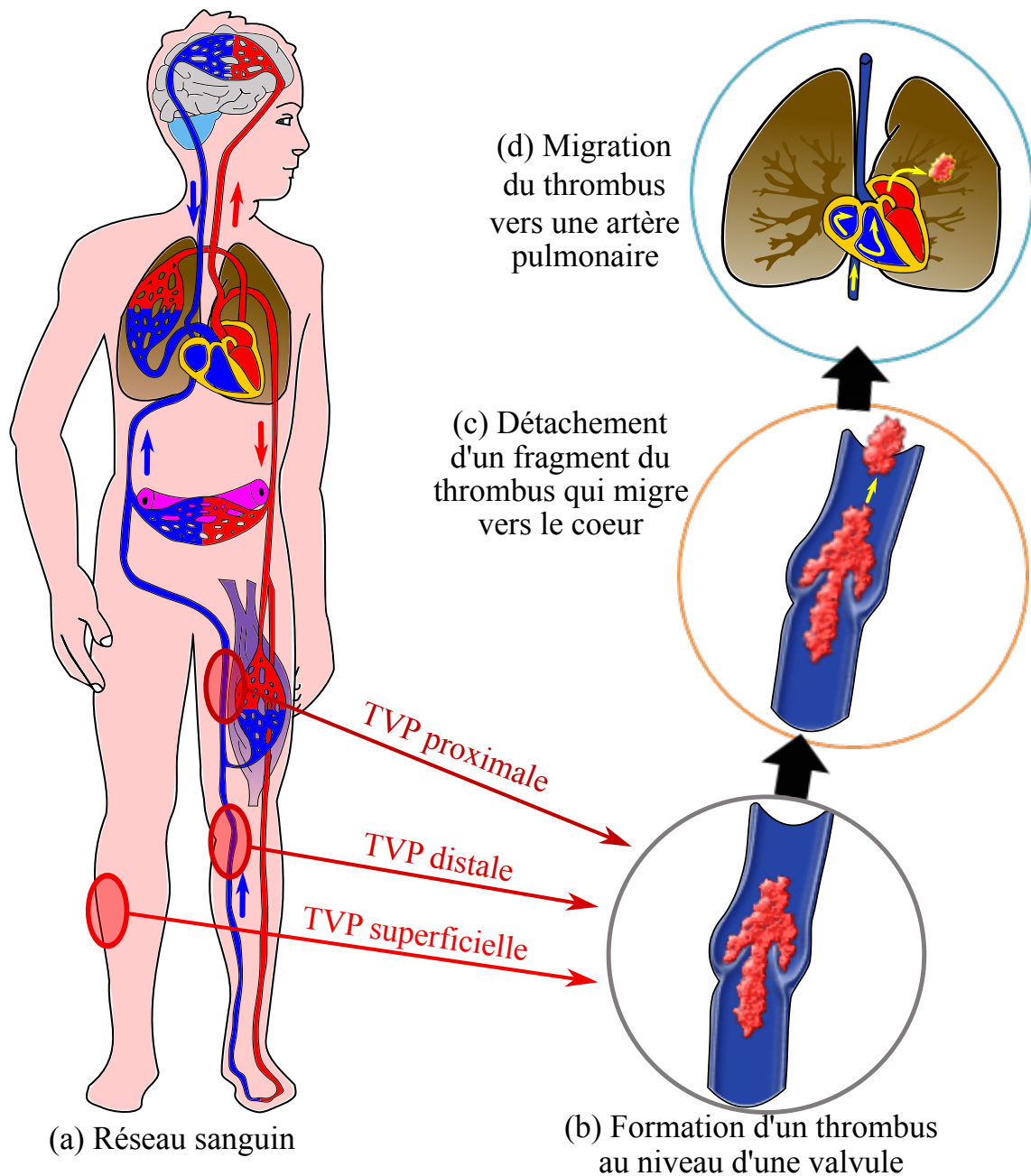


FIGURE 1.2 – De la thrombose veineuse à l'embolie pulmonaire (inspiré de [114, 135]).

1.1.4 Diagnostic et traitements

Afin d'éviter tout risque d'EP, il est nécessaire de détecter au plus tôt la thrombose veineuse [49]. En présence des symptômes (*cf.* section 1.1.3) et/ou en cas de risque important (*cf.* section 1.1.2), une échographie-doppler (principe expliqué en annexe A.2.2) est très souvent prescrite au patient afin de vérifier sa circulation sanguine. L'échographie-doppler permet de visualiser le réseau veineux, le flux sanguin ainsi que la compressibilité de la veine. D'autres techniques d'imagerie présentées en annexe A.2 peuvent être envisagées pour observer un thrombus mais l'échographie reste la solution privilégiée pour son efficacité, son coût, sa disponibilité et son caractère non invasif. La présence d'un thrombus se traduit par une veine incompressible (voir figure 1.3) et par une circulation sanguine perturbée voire absente. De plus, un test sanguin complémentaire peut-être réalisé afin de mesurer le taux de D-dimères (substance générée lors de la dégradation de la fibrine). Une faible quantité (inférieure à 500 microgrammes par litre de sang) de cette substance libérée par le thrombus couplée à une faible probabilité clinique (peu voire aucun facteur de risque) permet en principe d'exclure une TVP.

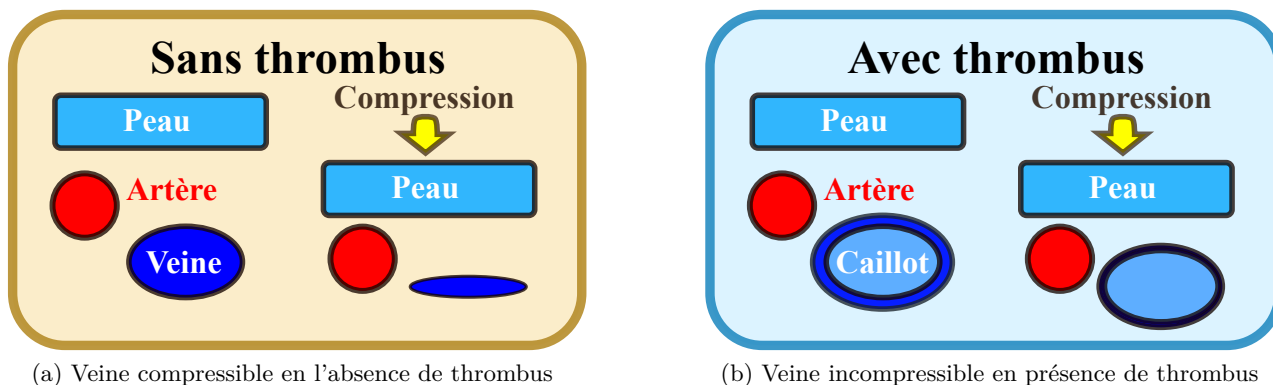


FIGURE 1.3 – Diagnostic d'une thrombose veineuse profonde en testant la compressibilité de la veine par échographie (en coupe transversale du thrombus).

En cas de TVP confirmée, des médicaments anticoagulants sont injectés par voie intraveineuse ou sous-cutanée en urgence afin de stabiliser le thrombus sanguin en empêchant son extension [49] : l'héparine de bas poids moléculaire, en injection quotidienne pendant cinq à sept jours est la plus utilisée avec un relais par traitement anticoagulant par voie orale (souvent les anti-vitamines K) afin d'éviter les récives. Ce traitement se poursuit sur une période allant de trois mois à plusieurs années, selon le patient et ses facteurs de risque. Le port quotidien de bas de contention est également recommandé dès le diagnostic et pendant au moins trois mois. Cela réduit l'œdème et prévient les complications (notamment le syndrome post-thrombotique).

Dans des cas extrêmes, une opération chirurgicale, l'angioplastie, peut être utilisée pour déboucher les veines profondes [77]. Cette intervention très délicate consiste à introduire dans la veine malade un cathéter muni à son extrémité d'un ballon gonflable. Le ballon est positionné au niveau du thrombus, puis gonflé pour permettre l'ouverture de la veine et améliorer le flux circulatoire. Les Stents ou endoprothèses sont des ressorts métalliques qui permettent de maintenir la veine ouverte après l'angioplastie. Lorsque le ballon est gonflé, le stent se détend à l'intérieur du vaisseau afin de lui éviter un nouveau rétrécissement. Toutefois, la finesse et la fragilité des parois veineuse (en comparaison avec les artères) rendent ce type d'opération beaucoup plus rare en veineux qu'en artériel.

1.1.5 Conclusion sur la problématique médicale de la thèse

Nous venons d'aborder les circonstances d'apparition d'un thrombus et les risques qu'il peut représenter s'il se forme de manière incongrue et/ou s'il ne se dissout pas correctement. La survenue des thromboses veineuses profondes est multifactorielle (immobilisation, cancer, prédisposition génétique, *etc.*) et leur origine est souvent difficilement identifiable. C'est cette raison qui nous a poussés à étudier la structure du thrombus pour déterminer si elle diffère suivant l'origine de la thrombose. Si c'est le cas, nous souhaitons pouvoir retrouver le principal facteur responsable de la thrombose à partir de la structure du thrombus ou prédire le risque d'embolie pulmonaire. Plusieurs techniques d'imagerie permettent d'observer les thrombi veineux comme la phlébographie, l'Imagerie par Résonance Magnétique (IRM), l'échographie ou l'élastographie. Ces deux dernières techniques exploitent les ondes ultrasonores pour générer des images de la région d'intérêt. Elles présentent donc l'avantage d'être indolores, non intrusives, rapides et peu coûteuses. L'échographie étant le plus souvent suffisante pour détecter une thrombose veineuse, l'usage de l'IRM, qui permettrait d'obtenir une meilleure qualité d'images, est difficilement envisageable au vu de sa disponibilité et de son coût.

1.2 Imagerie acoustique pour la caractérisation de tissus humains

Cette section est consacrée aux outils de diagnostic de la TVP via des ultrasons et à leur principe de fonctionnement. La section 1.2.1 décrit les phénomènes physiques se produisant lorsqu'une onde acoustique se propage. Nous aborderons d'abord le cas particulier du milieu élastique puis nous verrons l'extension des modèles présentés au corps humain. La section 1.2.2 présente les différents modes d'imagerie et le principe de fonctionnement d'un système échographique. Ces deux sections s'appuient sur les cours de M. Dauzat [40] et C. Cachard [53] en échographie médicale. Ensuite, la section 1.2.3 envisage une nouvelle façon d'étudier la structure d'un thrombus veineux en se penchant sur son élasticité (ou dureté).

1.2.1 Propagation acoustique

1.2.1.1 En milieu élastique

Un milieu élastique est un milieu qui reprend sa forme d'origine après avoir été déformé. Les ultrasons sont des vibrations mécaniques, variations locales de pression se propageant dans les milieux sans transport de matières. Ils se propagent dans un milieu élastique de densité volumique de masse ρ (ou masse volumique, en kg/m^3) selon deux modes de vibration définis en fonction des modules élastiques de compression K (en Pa) et de cisaillement μ (en Pa) :

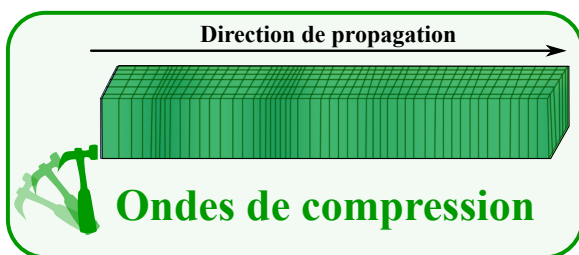
(a) une **onde longitudinale**, appelée aussi **onde de compression**, qui se déplace parallèlement à la direction de propagation à une vitesse définie par [69] :

$$c_l = \sqrt{\frac{3K + 4\mu}{3\rho}} \quad (1.1)$$

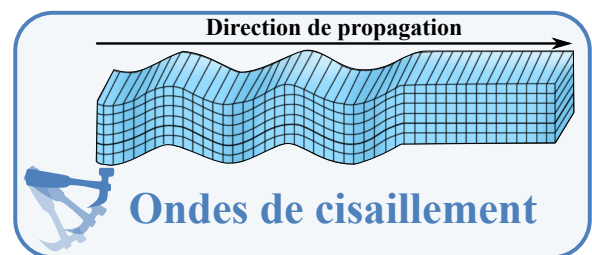
(b) une **onde transversale**, appelée aussi **onde de cisaillement**, qui se déplace perpendiculairement à la direction de propagation à une vitesse définie par [69] :

$$c_t = \sqrt{\frac{\mu}{\rho}} \quad (1.2)$$

Ces ondes sont représentées en figure 1.4. L'onde longitudinale se propage à une vitesse plus élevée ($\approx 1500 \text{ m/s}$) que l'onde transversale ($\approx 10 \text{ m/s}$) [69].



(a) Ondes de compression/dilatation (ou longitudinales) utilisées en échographie



(b) Ondes de cisaillement (ou transversales) utilisées en élastographie

FIGURE 1.4 – Types d'ondes acoustiques (inspiré de [127]).

En se propageant, les ultrasons subissent plusieurs phénomènes (voir l'annexe A.3 pour plus de détails) qui vont venir atténuer le signal. La perte la plus importante est consécutive à la transformation de l'énergie acoustique en énergie thermique. Lorsque les ultrasons rencontrent une interface (*i.e.* un changement de milieux), une partie de l'énergie est réfléchiée et l'autre partie continue de se propager dans le milieu. Sur des interfaces de très petites dimensions, l'énergie acoustique se disperse de façon homogène dans toutes les directions de l'espace (diffusion). Des phénomènes

de diffraction ou de réfraction peuvent se produire en fonction des conditions d'incidence sur les interfaces. Toutefois, l'échographie est principalement fondée sur la réflexion en ce qui concerne la délimitation des organes. L'atténuation englobe l'ensemble des pertes d'énergie dues à la propagation et décroît globalement de façon exponentielle avec la distance. Les systèmes d'échographie compensent en règle générale ces pertes en amplifiant le signal en fonction du temps de réception (réglage du gain).

Dans les liquides (comme le sang), le module d'élasticité μ est négligeable car l'amplitude de l'onde transversale s'atténue très rapidement. Par approximation, on considère que seule l'onde longitudinale peut se propager dans un liquide. Dans les tissus biologiques, principalement constitués d'eau, cette approximation est toujours utilisée (sauf en élastographie mais ce cas sera traité plus loin dans la section 1.2.2) et, lorsqu'une onde est mentionnée dans la suite de cette section, il s'agit de l'onde longitudinale. La vitesse de propagation d'une onde ultrasonore dans un liquide ou dans les tissus biologiques peut donc s'écrire :

$$\mu \ll K \Rightarrow c_l = \sqrt{\frac{K}{\rho}} \quad (1.3)$$

Le tableau 1.1 donne des ordres de grandeur de densités volumiques, de vitesses des ondes longitudinales et de leurs atténuations. On peut voir que les ondes se propagent beaucoup plus rapidement et sont moins atténuées dans l'eau que dans l'air. Lors du traitement des signaux échographiques, la densité volumique de masse ρ des tissus mous (composés majoritairement d'eau) est supposée constante et égale à $1000 \text{ kg} \cdot \text{m}^{-3}$, ce qui représente la densité ρ de l'eau. D'après le tableau 1.1, on voit également que l'air et les os atténuent fortement les ultrasons et représentent donc des obstacles lors des examens échographiques.

Milieux	Densité ρ (kg/m^3)	Célérité c_L ($\text{m} \cdot \text{s}^{-1}$)	Atténuation α (dB/cm^1) à 1 MHz
Air	1.20	344	12
Eau	1000	1480	0.0022
Sang	1060	1530	0.18
Graisse	920	1410 à 1470	0.6
Muscle	1070	1550 à 1640	1.2
Foie	1060	1550	0.9
Vaisseaux	1080	1530	
Tissus mous	≈ 1000	≈ 1540	
Os	1200 à 1810	2700 à 4100	20

Tableau 1.1 – Caractéristiques acoustiques de milieux de références et de tissus humains [53]. L'atténuation est définie en annexe A.3.

1.2.1.2 Dans le corps humain

Le corps humain est un milieu acoustique complexe composé de nombreux sous-milieux aux caractéristiques acoustiques propres à chacun. Les paramètres acoustiques de ces milieux, dont certains sont donnés dans le tableau 1.1, dépendent des conditions de mesure, de la température du milieu mais aussi d'autres facteurs moins quantifiables comme l'âge, l'état sain ou pathologique des organes. Les tissus humains sont donc des milieux non-linéaires. Dans de tels milieux, la vitesse de propagation de l'onde sonore ne sera plus constante comme dans un milieu linéaire. A chaque instant, les pics positifs de l'onde, en raison d'une compression tissulaire à ce niveau, auront une vitesse de propagation plus élevée que les pics négatifs où existe une relaxation tissulaire. Il en résulte une distorsion de l'onde qui évolue d'une forme sinusoïdale vers une forme en dent de scie (voir figure 1.5) et une modification du spectre avec la création de fréquences harmoniques.

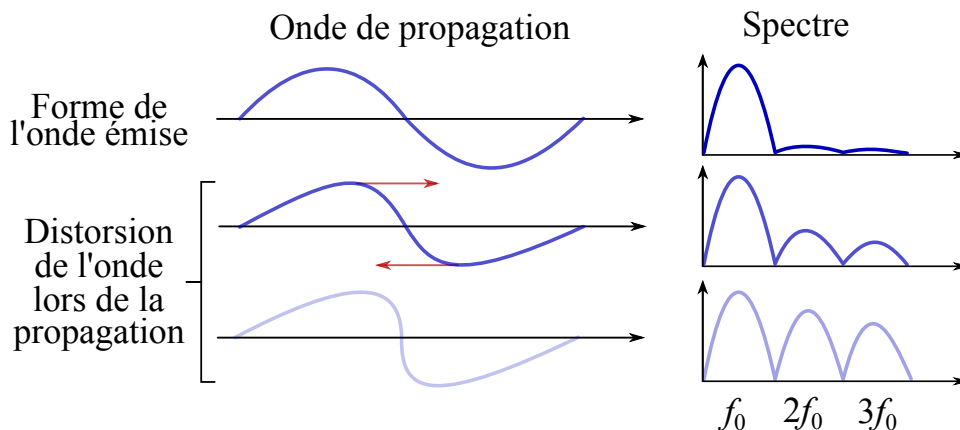


FIGURE 1.5 – Distorsion d’une onde acoustique de fréquence fondamentale f_0 dans un milieu non linéaire.

1.2.2 Echographie

1.2.2.1 Principe

Le principe de l’imagerie échographique est basé sur la détermination de l’amplitude et du retard d’un signal acoustique réfléchi par les interfaces d’un milieu pour en représenter sa structure. La sonde ultrasonore fonctionne en deux phases distinctes : l’émission et la réception. En phase d’émission, les transducteurs qui composent la sonde, excités par une impulsion électrique brève, génèrent un signal acoustique dans le milieu à explorer. En phase de réception, beaucoup plus longue, la sonde écoute les échos (ondes acoustiques réfléchies par les interfaces du milieu exploré) et les transforme en signaux électriques. Le signal d’émission transmis au système échographique sera envoyé seulement quand l’ensemble des échos sera parvenu à la sonde.

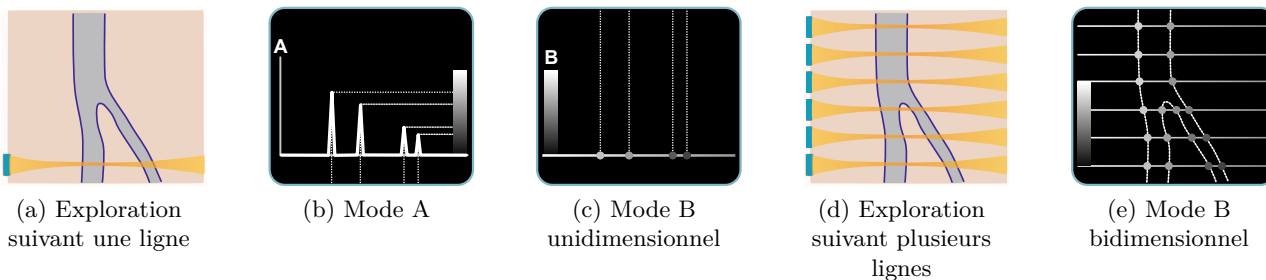


FIGURE 1.6 – Echographie en mode A ou B, unidimensionnelle ou bidimensionnelle [40].

Une fois le signal d’émission reçu, le système échographique va analyser l’amplitude du signal en fonction du temps (ou retard). Le retard traduit la profondeur de l’interface (tissu, organe, veine, etc.) et l’amplitude son échogénicité. Le mode A (Amplitude) est le mode le plus ancien et le plus simple. Il utilise l’émission d’une impulsion ultrasonore et la réception de son écho le long d’une seule ligne de propagation. L’information est généralement affichée sur un oscilloscope : les déflexions verticales (« pics ») qui s’affichent sur l’écran indiquent, par l’axe des abscisses, la profondeur de l’interface à l’origine de l’écho et, par l’axe des ordonnées, le coefficient de réflexion de l’interface (voir figure 1.6.b). Ce mode permet d’obtenir une mesure précise du retard et donc des distances séparant des structures échogènes mais n’est généralement pas utilisé pour les identifier.

Le mode le plus utilisé en imagerie médicale est le mode B (Brillance). Une image bidimensionnelle est formée par juxtaposition d’un grand nombre de lignes correspondant chacune à une échographie en mode A. Ces lignes sont obtenues soit par un déplacement manuel ou mécanique du transducteur, soit

en utilisant une barrette de transducteurs (le plus répandu actuellement) comme cela est schématisé en figure 1.6.d. Sur l'écran, chaque ligne correspond, en position et en direction, au trajet des ultrasons pour chaque position du transducteur. Le niveau de gris, appelé brillance, représente l'amplitude des échos (voir figure 1.6.c), libérant ainsi une dimension par rapport au mode A et permettant d'afficher une image (voir figure 1.6.e).

1.2.2.2 Imageries fondamentales et harmoniques

L'examen échographique (*cf.* figure 1.6) repose sur l'émission d'une onde ultrasonore, la propagation de cette onde dans le milieu ausculté et la réflexion partielle de l'onde à l'interface entre deux milieux à caractéristiques acoustiques différentes. En échographie classique, on considère que l'intensité acoustique et la variation de densité du milieu sont faibles, et que, par conséquent, les phénomènes physiques décrits en section 1.2.1.1 s'appliquent. Quand une onde ultrasonore, de fréquence fondamentale donnée f_0 (harmonique 0), traverse un tissu, la vitesse de propagation de cette onde est constante en l'absence de réponse non-linéaire du milieu et le signal réfléchi est un signal d'amplitude importante à la fréquence f_0 . L'exploitation de ce signal est utilisé afin de construire l'image échographique et ce mode d'imagerie est appelé mode fondamental (*cf.* figure 1.7.a).

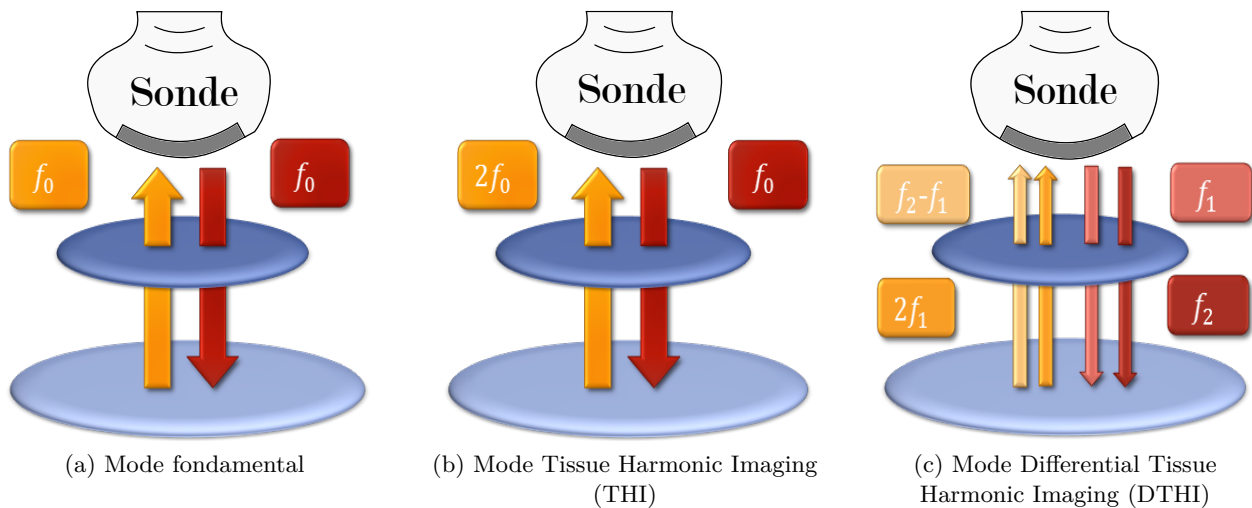


FIGURE 1.7 – Principe de fonctionnement des trois principaux modes d'échographie : fréquences des signaux émis et reçus.

En réalité, les tissus humains sont des milieux non-linéaires et l'onde ultrasonore est déformée au cours de sa propagation (*cf.* section 1.2.1.2). Le spectre de fréquence (*i.e.* la représentation fréquentielle du signal), initialement centré sur la fréquence fondamentale, s'enrichit peu à peu en fréquences harmoniques (fréquences multiples de la fondamentale). En mode fondamental, le système effectue un filtrage passe-bande afin de garder uniquement les échos à la fréquence fondamentale et éliminer les autres fréquences qui sont considérées comme du bruit (voir figure 1.7.a). Le bruit résiduel dégrade la qualité des images. Le mode harmonique, Tissue Harmonic Imaging (THI), corrige en grande partie ce défaut en supprimant la composante fondamentale de l'écho et en traitant uniquement les composantes harmoniques. En pratique, l'amplitude des harmoniques décroît fortement pour chaque multiple de la fréquence fondamentale donc les systèmes harmoniques s'intéressent uniquement aux échos de l'harmonique supérieure (*i.e.* $2f_0$, voir figure 1.7.b).

Un autre mode, développé par Toshiba, permet également d'améliorer sensiblement la qualité de l'image : Differential Tissue Harmonic Imaging (DTHI) [33, 85]. Dans ce mode, la sonde émet simultanément deux pulses de fréquences différentes f_1 et f_2 choisies telles que $f_1 < f_2 < 3f_1$. Les tissus humains étant des milieux non-linéaires, le récepteur de sonde va donc capter un signal composé des fréquences fondamentales f_1 et f_2 , de leur somme $f_1 + f_2$, de leur différence $f_1 - f_2$ et des harmoniques $2f_1$ et $2f_2$ (les autres harmoniques étant négligés car d'amplitudes trop faibles). Les composantes

fondamentales sont éliminées par une technique de soustraction d'impulsion (technique proche de l'inversion d'impulsion [33]). Les composantes des fréquences $f_1 + f_2$, $2f_2$ et des autres harmoniques seront, en pratique, filtrées en raison de la bande passante de la sonde. Au final, seules les composantes des fréquences $f_2 - f_1$ et $2f_1$ seront conservées et traitées (voir figure 1.7.c).

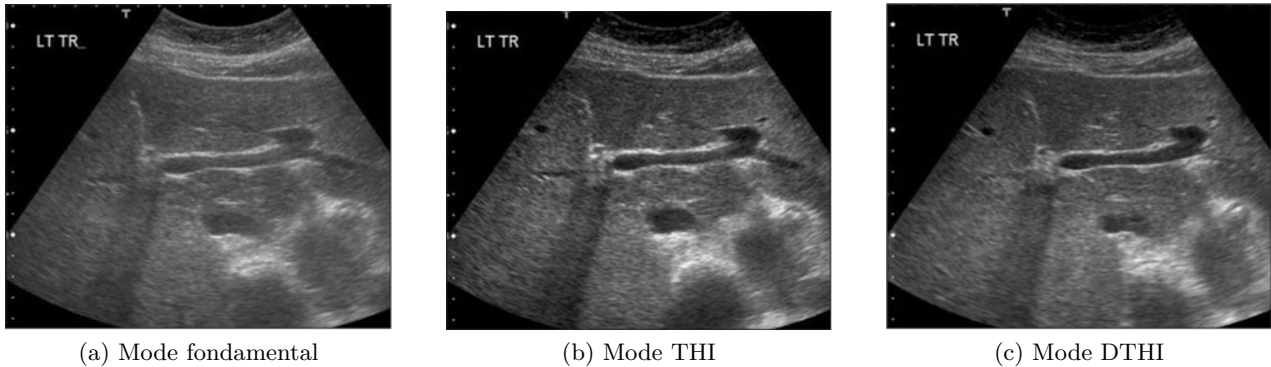


FIGURE 1.8 – Comparatif des trois modes d'échographie : les contours sont plus nets en THI et DTHI, l'image est moins atténuée en THI dans les profondeurs maximales [33].

Par rapport à l'imagerie harmonique classique où l'on traite uniquement la deuxième harmonique (voir figure 1.8.b), l'imagerie différentielle harmonique permet d'améliorer la pénétration des ondes, la définition des tissus et de leurs parois, sans altérer la résolution spatiale et de contraste. La figure 1.8) compare les trois techniques en imageries hépatiques : si le mode fondamental semble moins bon en terme de résolution en contraste, il est plus compliqué de départager les modes harmoniques sur cet exemple. D'après [33], les modes THI et DTHI sont beaucoup plus performants que le mode fondamental en imagerie hépatique. Les résultats montrent également que les images acquises avec le mode DTHI sont de meilleure qualité notamment au niveau de la définition des détails.

1.2.3 Elastographie

La dureté (ou l'élasticité) correspond à la capacité d'un milieu à se déformer lorsqu'on lui applique une contrainte mécanique (exprimée en kPa) : plus un milieu est dur, plus son élasticité augmente. Historiquement, la dureté des tissus, évaluée par palpation [49], est une mesure importante pour diagnostiquer une fibrose hépatique. En effet, en cas d'agression chronique du foie, le processus de cicatrisation va remplacer les cellules hépatiques endommagées par une cicatrice fibreuse. Plus le foie est dur, plus la fibrose est importante. Dans le cadre de cette thèse, l'objectif est d'analyser l'élasticité du thrombus qui, normalement, durci avec l'âge (car il est de plus en plus composé de fibrine). Néanmoins, le rôle de l'élastométrie dans le diagnostic de thrombose veineuse reste encore à définir. Les deux sections suivantes présentent les deux principales méthodes d'élastographie en s'appuyant sur les articles [60, 69].

1.2.3.1 Méthodes quasi-statiques

Lorsqu'une contrainte statique externe, σ , est appliquée sur la surface d'un solide (voir figure 1.9), il se produit une déformation, ϵ , qui dépend de la dureté du matériau E (module de Young). Lorsque la déformation est purement élastique (*i.e.* lorsque la déformation est réversible, par opposition à la déformation plastique qui est irréversible), la contrainte, σ (Pa), suit une loi linéaire, appelée loi de Hooke :

$$\sigma = \epsilon E \quad (1.4)$$

où E est le module de Young (Pa) et ϵ la déformation (adimensionnel).

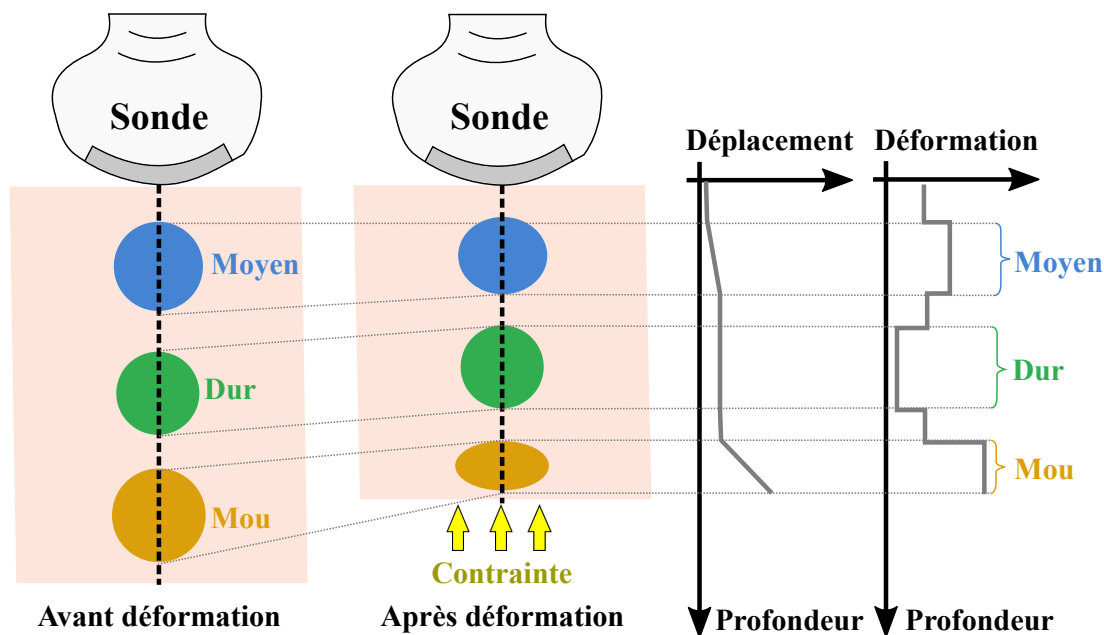


FIGURE 1.9 – Déformation élastique et déplacement des tissus biologiques sous l’action d’une contrainte externe en fonction de leur dureté.

Les méthodes quasi-statiques consistent à mesurer le déplacement et la déformation générée par corrélation des échographies sans et avec contrainte externe. En pratique, comme la contrainte appliquée est inconnue, seule une carte de la déformation ϵ (appelée aussi élastogramme) est affichée à l’écran. Ces méthodes ont l’avantage d’être simples à implémenter mais ne permettent pas d’estimer quantitativement le module de Young.

1.2.3.2 Méthodes dynamiques ou impulsionnelles

En section 1.2.1.1, on a vu que les tissus mous du corps humain peuvent être déformés sous l’action de deux types d’ondes mécaniques (illustrés en figure 1.4) : les ondes de compression utilisées en échographie et les ondes de cisaillement qui sont négligées en échographie due à leur forte atténuation. Toutefois, à plus basses fréquences (10 Hz à 2000 Hz [60]), les ondes de cisaillement sont moins atténuées et leur vitesse est directement liée au module de Young par la formule :

$$E = 3\rho c^2 \quad (1.5)$$

où ρ représente la masse volumique du tissu (kg/m^3) et c la vitesse des ondes de cisaillement (m/s). La masse volumique varie suivant le type de tissu ($\approx 920 kg/m^3$ pour la graisse, $\approx 1060 kg/m^3$ pour le sang et le foie sain, $\approx 1070 kg/m^3$ pour les muscles et entre 1200 et 1810 kg/m^3 pour les os [53]) mais elle est supposée constante ($1000 kg/m^3$) dans les calculs de la vitesse.

Les différentes méthodes dynamiques ne mesurent donc pas directement la dureté, mais la vitesse de propagation des ondes de cisaillement (shear waves) dans les tissus. Cette dernière est d’autant plus importante que le milieu est dur. Les systèmes émettent une impulsion soit mécanique, soit ultrasonore pour générer des ondes de cisaillement vers une région d’intérêt et, en parallèle, émettent des ultrasons (à une fréquence plus élevée) pour suivre la propagation des ondes de cisaillement dans les tissus. La figure 1.10 montre comment sont générées les ondes de cisaillement à partir d’impulsion ultrasonore. Le traitement des échos (par des méthodes d’intercorrélations complexes) permet ensuite de créer une carte de dureté ou de vitesse (appelée élastographie) mais aussi d’estimer la dureté (ou vitesse) moyenne d’une zone spécifiée par l’utilisateur (appelée élastométrie).

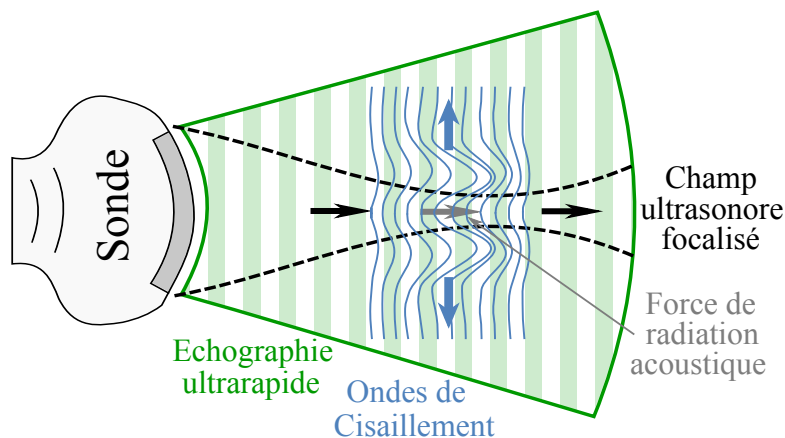


FIGURE 1.10 – Principe de l'élastographie dynamique par force de radiation : le champ ultrasonore focalisé (flèches noires) interagit avec le milieu en créant une force de radiation (flèche grise). Cette dernière va générer des ondes de cisaillement (flèches bleues) dans la direction perpendiculaire à l'axe du champ ultrasonore. La propagation de ces ondes est suivie au cours du temps par échographie ultrasonore (inspiré de [89]).

Ces méthodes dynamiques possèdent l'avantage de pouvoir produire des images haute-résolution et quantitatives du module de Young mais elles nécessitent des systèmes plus complexes (que les méthodes quasi-statiques) capables de générer des ondes de cisaillement et de mesurer les petits déplacements induits par ces ondes de cisaillement. De nombreux constructeurs proposent des systèmes capables d'estimer la dureté d'un organe ou de tissus : Echosens, Hitachi, Siemens, General Electric, Philips, Toshiba, Supersonic, Ultrasonix et Esaote. Le CHU de Brest et plus particulièrement l'Unité d'Echodoppler et de Médecine Vasculaire, dispose du système de Toshiba Aplio 500 [160] qui a l'avantage de proposer les deux types de méthodes, quasi-statiques et dynamiques, pour cartographier la dureté du thrombus.

1.2.4 Conclusion sur l'imagerie acoustique en médecine

Dans cette section, nous avons remarqué que deux types d'ondes acoustiques (compression et cisaillement) peuvent se propager dans le corps humain et l'analyse de ces deux types d'ondes permet de générer deux modes d'imagerie distincts : l'échographie et l'élastographie. L'échographie permet de visualiser une région d'intérêt (organes, tissus, vaisseaux sanguins, os, *etc.*) en analysant la réflexion des échos sur les différents interfaces qu'ils vont rencontrer dans la région inspectée. L'élastographie s'intéresse à la vitesse de propagation des ondes de cisaillement (beaucoup plus lentes que celles de compression) qui est proportionnelle à la dureté du milieu qu'elles traversent. Elle permet donc de cartographier la dureté d'une région d'intérêt et notamment de détecter des tissus anormalement durs (fibrose du foie par exemple). Dans la section suivante, nous allons voir comment nous avons appliqué ces techniques d'imagerie pour construire notre base de données de thrombi veineux.

1.3 Collecte des images échographiques et élastographiques

La section 1.1 a défini les notions de thrombus, thrombose et d'embolie pulmonaires puis la section 1.2 a introduit deux techniques d'imageries acoustiques (échographie et élastographie) qui permettent de visualiser des thrombi veineux. Cette section présente tout d'abord le système que nous utilisons dans notre projet pour collecter les données : le Toshiba Aplio 500. Puis, la sous-section 1.3.2 détaille le protocole d'acquisition des données et son évolution au cours de la thèse. Enfin, la sous-section 1.3.3 introduit une autre base de données contenant des images de glandes salivaires acquises dans des conditions similaires.

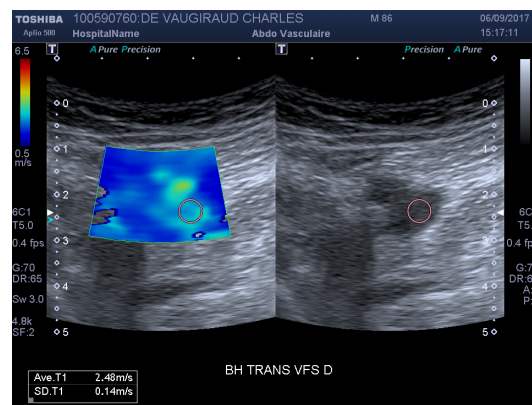
1.3.1 Système d'acquisition

1.3.1.1 Echographe Toshiba Aplio 500

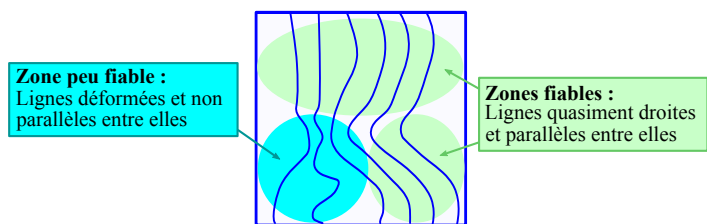
Dans notre procédure expérimentale, nos données sont acquises uniquement avec le système Aplio 500 de Toshiba (voir figure 1.11.a). Ce système dispose des trois modes d'échographie décrits en section 1.2.2.2 (fondamental, THI et DTHI) et des deux méthodes d'élastographie introduites en section 1.2.3 (quasi-statique et dynamique). Les sondes utilisées, qui jouent un rôle clé dans la qualité d'images, sont présentées en section 1.3.1.2. Lors du diagnostic de la thrombose veineuse, l'expert médical commence par examiner le réseau veineux du patient en ajustant au mieux les paramètres échographiques (*cf.* section 1.3.1.3). N'ayant pas accès aux données brutes, l'objectif est de modifier au minimum ces paramètres dans un souci de standardiser nos acquisitions et ainsi de limiter le biais pouvant être induit en comparant des images acquises par ces changements de paramètres lors de l'analyse des images. Toutefois, l'anatomie, la profondeur de la veine, la présence d'œdème sont autant de facteurs pouvant impacter la qualité de l'image et nécessiter une adaptation des paramètres (comme le gain général). Ensuite, l'expert médical effectue le test de compressibilité en mode échographique (*cf.* section 1.1.4) afin de détecter la présence de thrombus. En présence d'un thrombus, il positionne la sonde au niveau de la tête du thrombus et passe en mode d'élastographie dynamique.



(a) Système Toshiba Aplio 500 [160]



(b) Elastographie (partie gauche), échographie (partie droite) et élastométrie : moyenne (Ave.T1) et écart-type (SD.T1) des mesures de vitesse des ondes de cisaillement (en bas)



(c) Fiabilité des ondes de cisaillement pour estimer la qualité de la mesure [68]

FIGURE 1.11 – Echographie et Elastographie avec le système Toshiba Aplio 500.

1.3.1.2 Sondes

Les sondes sont des outils de hautes technologies qui jouent un rôle clé dans la qualité de l'imagerie médicale. Les caractéristiques principales qui interviennent dans le choix d'une sonde sont sa fréquence de résonance et sa bande passante. La sonde doit être adaptée au milieu à explorer et notamment à la profondeur de la zone cible. Diminuer la fréquence permet d'accroître la profondeur d'exploration mais diminue la résolution axiale (inverse de la distance minimale entre deux structures réfléchissantes distinguables par la sonde) de l'image échographique. Les résolutions latérale (direction parallèle à la largeur des transducteurs) et azimutale (direction parallèle à la longueur des transducteurs) sont déterminantes dans la qualité de l'image. Elles dépendent de la géométrie du transducteur, de la fréquence utilisée et de la focalisation du faisceau.

Sur le marché, il existe trois principaux types de sonde (voir figure 1.13) [40] :

- les **sondes linéaires** construisant une image rectangulaire par balayage linéaire avec des lignes d'exploration sont parallèles entre elles ;
- les **sondes convexes** formant une image en secteur de cercle par balayage linéaire avec des lignes d'exploration divergentes (appelé aussi balayage semi-sectoriel) et une zone de contact peau/sonde importante ;
- les **sondes sectorielles** formant également une image en secteur de cercle par balayage sectoriel avec des lignes d'exploration divergentes mais avec une zone de contact peau/sonde étroite.

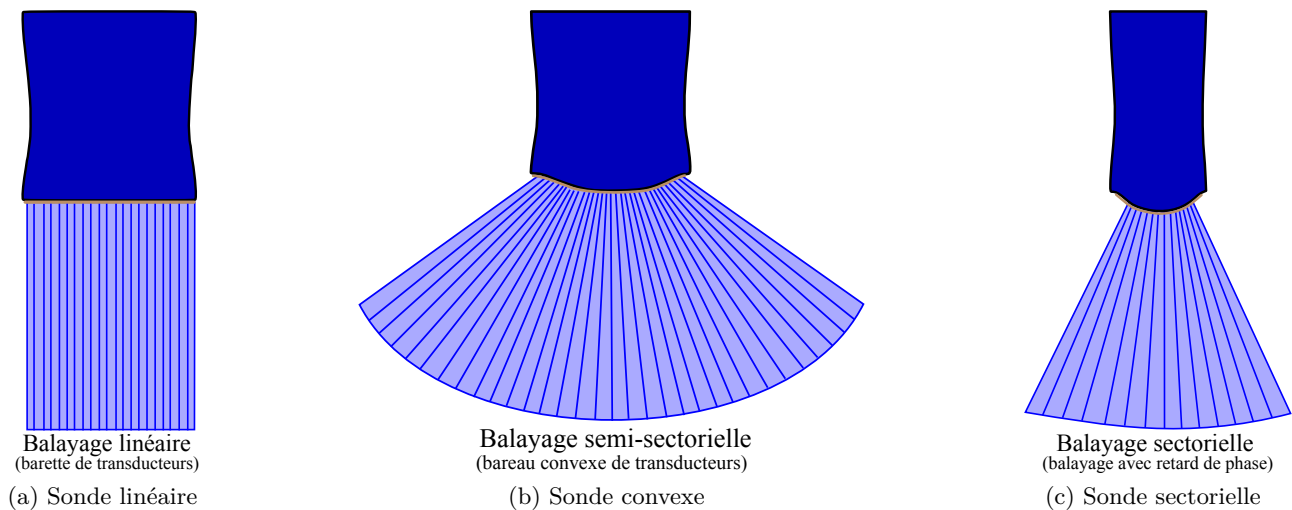


FIGURE 1.13 – Principaux types de sondes [40].

Pour notre base de données (BDD), deux sondes de Toshiba ont été utilisées [157] :

1. La sonde **14L5** (PVT-1005BT) qui est une sonde linéaire, de bande passante plus large [7 – 14] *MHz*.
2. La sonde **6C1** (PVT-375BT) qui est une sonde convexe, de bande passante [1 – 6] *MHz* destinée à des examens abdominaux (ce qui permet d'atteindre le réseau veineux profond des jambes même en cas de gonflement des jambes) ;

Le choix de la sonde est indiqué directement sur l'écran (6C1 ou 14L5) ou dans les métadonnées de l'acquisition.

1.3.1.3 Paramètres d'acquisition

Comme évoquer précédemment, lors de l'examen, l'opérateur médical peut ajuster certains paramètres d'acquisition afin de visualiser au mieux le thrombus veineux. Les paramètres sélectionnés lors de l'acquisition sont visibles directement sur l'écran d'affichage (voir figure 1.12). La suite de cette section passe en revue la liste des principaux paramètres ajustables en s'appuyant sur la figure 1.12 et sur les livres [40, 129] :

- **Focalisation :**

La sonde est composée de plusieurs transducteurs afin de mieux maîtriser le faisceau ultrasonore émis. La focalisation électronique permet ainsi à l'utilisateur d'améliorer la résolution latérale dans la zone focale. Cette zone est indiquée par la tête d'une flèche (*i.e.* un triangle, situé au niveau du champ 3 de la figure 1.12). La partie droite de l'écran affiche l'échographie en mode B : l'échelle de gris (champ 2 de la figure 1.12) traduit l'échogénicité des interfaces rencontrées par le faisceau ultrasonore.

- **Région d'intérêt :**

La ROI de forme elliptique est dessinée par l'opérateur lors de chaque acquisition et englobe une partie (comme sur la figure 1.12) ou la totalité (comme sur la figure 1.11.b) du thrombus. En mode élastographie, elle permet de définir la zone dans laquelle on souhaite estimer la moyenne et l'écart type (champs 16 et 17 de la figure 1.12) des vitesses des ondes de cisaillement. La valeur moyenne est aussi appelée élastométrie. Il est également possible de dessiner plusieurs ROIs sur une même image afin d'obtenir les élastométries dans les différentes ROIs.

- **Mode et fréquence de l'échographie :**

L'opérateur peut adapter la fréquence d'émission en fonction de la profondeur d'exploration voulue et opter parmi les trois modes d'échographie décrits en section 1.2.2.2 : fondamental, THI (indiqué par un « T » sur écran) et DTHI (indiqué par « diffT »). La fréquence d'émission est inscrite dans les métadonnées du fichier DICOM. En THI et DTHI, la fréquence de réception est différente de la fréquence d'émission. Cette dernière est indiquée dans le champ 6 (figure 1.12) juste après le mode (rien, « T » ou « diffT »).

- **Cadence d'image :**

La résolution temporelle dépend du nombre d'images acquises par secondes. En cas de cadence d'imagerie supérieur à 16 images par secondes, l'opérateur à l'impression d'effectuer une exploration en « temps réel ». La cadence d'imagerie est inversement proportionnelle à la profondeur d'exploration et au nombre de lignes qui composent l'image. Dans notre cas, la cadence est de 0.4 images par seconde (champ 7 de la figure 1.12) et la profondeur d'exploration est d'environ 5 *cm* (variable suivant les patients).

- **Gain :**

Les ondes acoustiques sont atténuées lorsqu'elles se propagent dans les tissus biologiques. Le réglage du gain général (champ 8 de la figure 1.12) permet de compenser ses atténuations. Une augmentation du gain entraîne une augmentation de l'amplitude des échos en réception. L'image devient plus lumineuse, mais cette amélioration a un coût puisqu'elle va baisser la résolution spatiale et introduire du bruit. Toutefois, les interfaces peu échogènes peuvent ne pas être visibles si le gain est trop faible. La fonction « Qscan » du système Toshiba permet d'ajuster automatiquement le gain (les traitements utilisés pour optimiser le gain ne sont pas décrits par le fabricant) qui, par conséquent, peut varier suivant les images de notre base de données. Il est également possible d'ajuster le gain à des profondeurs spécifiques (Time Compensation Gain).

- **Dynamique de l'image :**

En mode B, l'échelle de gris traduit l'échogénéité des interfaces rencontrées par les ondes acoustiques. Le nombre de niveau gris détermine la résolution en contraste. Elle dépend de l'étendue de la gamme d'intensités acoustiques que le système d'acquisition est capable de détecter et de restituer (sensibilité de la sonde et qualité d'amplification). La dynamique des images est indiquée par le champ 9 (figure 1.12) sur l'écran et est fixée à 65 *dB* pour la grande majorité des acquisitions de notre base de données (55 *dB* pour quelques images d'un patient). La dynamique est un paramètre important si l'on souhaite comparer les informations de contraste d'images de plusieurs patients.

- **Indices de qualité :**

Le système de Toshiba permet à l'utilisateur d'améliorer la qualité de l'échographie en jouant sur deux indices (1.12 champ 10) [121] : « Precision + » (P) et « Aplipure » (A). Le premier indice P permet d'améliorer la résolution en contraste en prenant compte les multi-trajets. En effet, l'image échographique en mode B est, de base, formée ligne par ligne et le niveau gris de chaque pixel correspond à l'amplitude de l'écho qui est proportionnelle à la réflectivité des interfaces. Avec l'outil P, l'image ne considère pas seulement les trajets directs (i.e. en ligne droite) mais également les trajets multiples dus aux variations de célérité des tissus biologiques. Quant au second indice A, il utilise la technique d'imagerie de combinaison ou *compounding*. Le principe de cette technique est de moyennner des images acquises à des incidences et à des fréquences différentes. En pratique, en ce qui concerne la combinaison spatiale, le formateur de faisceau balaie la tranche de section avec des angles progressivement croissants, puis décroissants et l'image est reconstruite à partir de 5 à 9 balayages spatiaux. L'imagerie de combinaison permet de diminuer le bruit et les artéfacts tout en rendant les contours plus nets et l'image plus lisse.

- **En élastographie :**

L'élastographie permet d'obtenir une carte de dureté (en *kPa*) ou de vitesses des ondes de cisaillement (en *m/s*) dans une région d'intérêt (champs 11 de la figure 1.12). Le principe de cette technique a été développé à la section 1.2.3 et est notamment basé sur le module de Young. Durant une acquisition, la sonde va émettre des ondes de cisaillement dont elle va simultanément suivre l'évolution grâce aux ondes de compression (échographie classique). Dans notre étude, la fréquence des ondes de cisaillement et la fréquence de répétition des impulsions (champs 13 et 14 de la figure 1.12) dépendent de la sonde utilisée. L'opérateur peut ajuster le paramètre nommé Spatial Format (figure 1.12 champ 15) qui joue sur le filtre spatial de lissage des vitesses. Comme nous l'avons vu en figure 1.11.c, l'utilisateur peut vérifier qualitativement la qualité de l'élastographie en affichant les lignes de propagation des ondes de cisaillement dans la région d'intérêt.

1.3.2 Evolution du protocole d'acquisition des données de thrombi

Cette section décrit l'évolution du protocole expérimental utilisé pour la collecte des données échographiques et élastographiques. L'objectif du protocole est de trouver le meilleur compromis afin d'obtenir la meilleure qualité d'image en changeant le moins possible les conditions d'acquisition (localisation, fréquence, mode d'échographie, gain, *etc.*).

1.3.2.1 Protocole 1 : pré-mise à jour

Le premier protocole d'acquisition consiste à observer la tête du thrombus de façon transversale avec la sonde convexe (6C1). Deux opérateurs se relayent et effectuent une dizaine d'acquisitions chacun. Le tableau 1.2 liste l'ensemble des patients examinés et classés en fonction de leur épidémiologie (cancer, « Immob. » ou « Im. » pour une immobilisation, « Chirur. » ou « Ch. » pour une chirurgie, « Gross. » pour une grossesse, « Idiop. » ou « Id. » pour une thrombose idiopathique) et de la survenue d'une EP. Dans ce tableau, on y trouve d'autres informations sur les différents patients comme son âge, son sexe, la localisation de la thrombose (gauche, droite ou les deux), le facteur déclencheur connu et la présence d'antécédents personnels.

Classes	ID	Age	Sexe	TVP	Déclencheur	Ant.	Mois	Op1	Op2	Echo	Gain	
Cancer	013	75	F	B			0	10	11	THI 5	75	
	015	85	M	B			0	10	10	THI 5	75	
Immob.	002	54	F	G	Voiture (7h)	Oui	0	11	11	DTHI 5	80	
							3	11	10	THI 5	75	
	018	90	F	G	Immobilisation		0	11	10	THI 5	75	
Im. EP	014	56	M	D	Hospitalisation		0	10	9	DTHI 5	83	
Chirur.	005	94	F	D	Chirurgie		0	10	11	THI 5	64	
Ch. EP	009	72	M	G	Chirurgie		0	10	10	THI 5	75	
Gross.	017	30	F	G	Grossesse		0	10	10	THI 5	75	
							3	11	13	THI 5	75	
	019	28	F	G	Grossesse		0	11	9	F 2 / THI 5	77 / 75	
Idiop.	001	75	F	D			0	10	10	THI 5	75 / 71	
	003	80	F	D			0	9		THI 5	60	
	004	66	M	G			0	10	11	THI 5	75 / 74	
	006	40	M	G			0	11	10	THI 5	67	
							3	10	11	DTHI 5	71	
	008	80	M	G			0	10	21	THI 5	75	
011	44	M	D			Oui	0	10	11	THI 5	75 / 64	
Id. EP	007	46	M	G			0	10	10	THI 5	60	
	010	66	F	G			0	10	10	THI 5	75	
							3	10	10	THI 5	5	
	012	70	F	G			Oui	0	12	15	THI 5	68
	016	99	F	G				0	10	10	THI 5	75 / 72

Tableau 1.2 – Liste des acquisitions par patient obtenues avec le protocole 1 : classes, identifiant (ID), âge, sexe (Féminin/Masculin), TVP (Gauche, Droite ou Bilatérale), facteur déclencheur, antécédent (Ant.), mois d'observation (M), nombre d'images par opérateur (Op1 et Op2), mode et fréquence de l'échographie (Echo.), et gain.

Pour chaque patient, les experts médicaux conservent au maximum les mêmes paramètres d'acquisition. Toutefois, certains paramètres comme le gain et le mode d'échographie (fondamental, THI et DTHI) sont adaptés pour observer le thrombus. Le tableau 1.2 donne le nombre d'acquisitions effectuées par chaque opérateur, le mode d'échographie choisi et le gain. On peut voir que la majorité des examens sont réalisés avec l'imagerie harmonique (THI ou DTHI) et que le mode fondamental a été utilisé une fois pour le patient 19. En outre, un examen de contrôle est, si possible, réalisé 3 mois après le premier contrôle. Lorsque le thrombus est toujours présent, une série d'acquisitions est de nouveau menée et est indiquée par la colonne M (Mois 0 ou 3) dans le tableau 1.2.

1.3.2.2 Protocole 2 : post-mise à jour

Depuis juin 2016, le système Toshiba Aplio 500 a été mis à jour et une nouvelle sonde linéaire (14L5) a été mise à disposition par Toshiba. Cette seconde sonde fonctionne à des fréquences plus élevées (entre 7 et 14 Mhz contre 1 et 6 MHz pour la convexe) ce qui devrait générer des images de meilleures résolutions. La mise à jour du système inclut de nouveaux prétraitements des images échographiques mais surtout une amélioration du mode élastométrie avec notamment une meilleure estimation des vitesses des ondes de cisaillement (voir figure 1.14).

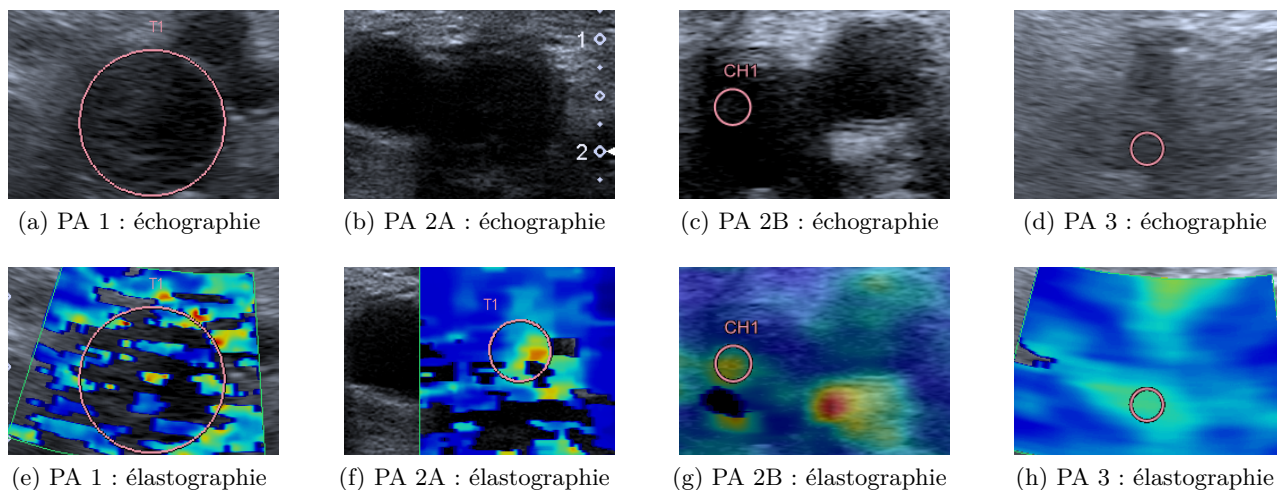


FIGURE 1.14 – Comparaison des échographies et des élastographies en fonction du Protocole d’Acquisition (PA).

La principale limite de la base obtenue avec le protocole 1 réside dans la qualité des élastographies. En effet, celles-ci comportent de nombreuses zones d’ombre (zone sans valeur de vitesse) au sein d’une même image ; et les images varient grandement d’une acquisition à l’autre. L’apport de la sonde linéaire et de la mise à jour devrait donc améliorer la qualité des élastographies. De plus, pour minimiser la variabilité intra- et inter-opérateurs et ainsi rendre nos acquisitions plus stables, l’équipe médicale a mis au point un bras articulé qui tient la sonde contre la jambe du patient. L’utilisation de ce bras permet d’avoir une pression constante de la sonde sur le patient entre plusieurs opérateurs et pour différents patients.

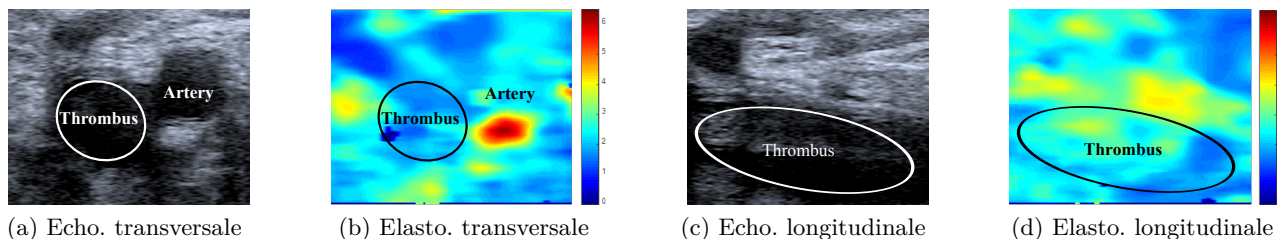


FIGURE 1.15 – Echographies et élastographies en mode transversal et longitudinal.

Lors du protocole précédent, le thrombus est observé uniquement dans le sens transversal, or une observation longitudinale peut être également une source d’information pertinente pour analyser la structure du thrombus (voir figure 1.15). De plus, on a vu précédemment que les images exportées sont des captures d’écran et les annotations inscrites par l’expert sur l’image (ellipse autour du thrombus par exemple) ne peuvent pas être supprimées *a posteriori*. Durant une série d’acquisitions, des captures d’écrans sont donc effectuées sans et avec les annotations. En mode élastographie, les lignes de propagation sont contrôlées en vérifiant que les lignes soient bien parallèles entre elles ; la carte de propagation est également exportée. Pour chaque patient, le nombre d’acquisitions par opérateur a été réduit de dix à deux ou trois afin que le temps de l’examen médical ne soit pas allongé.

Classes	Epidémiologie								Echographie				Elasto.	
	ID	Age	Sexe	TVP	Déclencheur	Gén.	Tab.	Mois	Mode	Gain	T	L	T	L
Cancer	012	56	M	FCD	Chirurgie	Fam.	Oui	0	DTHI 14	73	1	3	1	3
	013	81	F	FCG				0	DTHI 14	88	1	5	1	7
	015	52	M	FCD				0	DTHI 10	79	1	2	2	3
	016	84	M	FCG	Immobilisation			0	DTHI 10	96	1	1	1	3
	021	61	M	FCG				0	DTHI 14	93	1	4	5	3
Ca. EP	006	84	M	FSG				0	DTHI 10	80			2	7
	011	93	F	FCD				0	DTHI 14	80	1		2	6
	018	95	F	FCG				0	DTHI 14	79	1	3	2	3
	022	91	F	FSD				0	DTHI 10	78	1	5	2	6
Immob.	008	79	F	FCG	Allitement			0	DTHI 14	77	1		1	5
Im. EP	004	65	M	FCD	Voiture (16h)	Ant.		0	DTHI 10	67	1		3	6
		51	M	PD	Plâtre			0	DTHI 14	78	3	3	1	2
Gé. EP	019	91	F	FCD		Ant.		0	DTHI 14	73	2	2	1	1
		37	M	ICG		FVL	Oui	0	DTHI 10	78	1		2	4
Idiop.								1	DTHI 14	87	3	3	3	8
	001	88	F	PD				0	DTHI 14	73	2	4	2	5
	002	65	M	FCDG				0	DTHI 10	84	1	1	1	4
	009	84	M	FCG				0	DTHI 10	78	1	1	1	4
	017	86	F	FCG					0	THI 11	85	1	3	1
1									DTHI 10	78	3	4	3	3
Id. EP	005	65	F	FSD				0	DTHI 10	78	1	1	2	6
	007	57	M	PG				0	DTHI 10	82	0		2	3
	010	57	M	FCG				0	DTHI 14	84	1		3	3
	014	66	M	FSG		Ant.		0	DTHI 14	100	1		2	5
	023	78	F	PD			Oui	0	DTHI 14	82	1	2	1	2

Tableau 1.3 – Liste des acquisitions par patient obtenu avec le protocole 2 : classe, identifiant (ID), âge, sexe (féminin/masculin), TVP (Commune, Droite, Fémorale, Gauche, Iliaque et/ou Poplitée), facteur déclencheur, prédispositions génétiques (antécédents familiaux « Fam. », antécédents personnel « Ant. » ou Facteur V Leiden « FVL »), tabac (Tab.), mois d'examen (M), mode et fréquence de l'échographie (Echo.), gain, nombre d'échographies et d'élastographies transversales (T) et longitudinales (L).

Le tableau 1.3 référence la liste des patients examinés et les conditions de survenue de la thrombose. Par rapport au protocole précédent, la prise de tabac a été ajoutée à l'épidémiologie ainsi que la présence de prédisposition génétique (« Gé. ») : mutation génétique (facteur V Leiden, FVL), antécédents personnels (« Ant. ») ou familiaux (« Fam. »). Il donne aussi le nombre d'échographies et d'élastographies effectuées pour chaque patient en mode transversal et en mode longitudinal. A trois mois, les thrombi étant souvent dissouts par le traitement anticoagulant, l'examen de contrôle est à présent effectué un mois après la découverte de la thrombose (si possible). On peut constater que la grande majorité des échographies ont été réalisées avec le mode DTHI mais que la fréquence alterne entre 10 MHz (DTHI 10) et 14 MHz (DTHI 14) suivant les patients. Si le gain varie d'un patient à l'autre, les autres paramètres (dynamique, index de qualité, fréquences de l'élastographie, spatial format) restent stables pour une fréquence d'échographie donnée comme indiqués dans le tableau 1.4.

Paramètres Protocoles	Echographie					Elastographie		
	Mode	f_i	D	A	P	f_i	f_r	SF
2A	DTHI	10 MHz	65	1	5	10 MHz	6 kHz	1
2B	DTHI	14 MHz	65	2	7	10 MHz	6 kHz	3

Tableau 1.4 – Paramètres d’acquisition du protocole 2 : mode, fréquence d’échographie f_t , dynamique et index de qualité A et P pour l’échographie, fréquence de la porteuse f_i , fréquence de répétition f_r et spatial format pour l’élastographie.

1.3.2.3 Protocole 3 : sonde convexe

Au fur et à mesure des acquisitions, l’équipe médicale a observé, malgré la mise à jour du système et l’usage du bras, des variations importantes des cartes d’élastographie pour un même patient et un même opérateur. Afin d’écartier l’impact de la pression artérielle sur l’élastographie, des essais réalisés sur des personnes décédées ont montré les mêmes problèmes de stabilité. Des mesures menées avec la sonde convexe de plus basse fréquence (utilisée dans le protocole 1), ont donné des résultats beaucoup plus stables. La profondeur nécessaire pour atteindre le réseau veineux est donc probablement trop importante pour pouvoir utiliser la sonde linéaire. Par conséquent, la troisième base de données a été réalisée à l’aide de cette sonde convexe.

Le tableau 1.5 présente les données collectées avec le protocole 3 : épidémiologie des patients, nombre d’acquisitions par type d’imagerie et par coupe (transversale ou longitudinale). On peut noter que les échographies ont été réalisées avec le mode THI ou le mode fondamental (FGS). La trentaine de patients est triée en fonction des facteurs de risque et de la survenue d’une EP. Lors des simulations présentées au chapitre 5, l’objectif est de trouver des corrélations entre l’épidémiologie des patients et les résultats de classification. La figure 1.14 montre l’évolution de l’élastographie en fonction du protocole utilisé. On peut voir que la mise à jour du système a amélioré considérablement la qualité des images. La différence entre les protocoles 2 et 3 est moins visible sur cette figure car elle réside davantage dans la stabilité des mesures : avec le protocole 3, il y a moins de variations au niveau des élastographies d’un même patient, d’une même zone et réalisées par le même expert médical.

Classes	Epidémiologie								Echographie				Elasto.	
	ID	Age	Sexe	TVP	Déclencheur	Gén.	Tab.	Mois	Mode	Gain	T	L	T	L
Cancer	036	57	M	FCG			OUI	0	FGS 5	75	2	7	2	7
	057	54	F	FSD			OUI	0	THI 5	66	1	2	1	2
	035	70	M	FCG				0	THI 5	70	2	6	2	9
	025	92	M	FCG	Immobilisation			0	FGS 5	75	1	7	1	7
	037	68	F	FCG	Allitement			0	FGS 5	75	2	3	4	7
	058	94	F	FCG	Allitement			0	THI 5	67	1	3	1	2
Ca. EP	030	68	F	FSD	Allitement			0	FGS 5	71	3	3	3	3
	038	82	M	FSD				0	FGS 5	75	1	7	2	7
	051	68	M	PG			OUI	0	THI 5	76	2	4	2	4
	055	60	M	FSD		Hor.	OUI	0	THI 5	75	2	4	2	5
	033	80	F	PBD	Chirurgie			0	FGS 5	75	4	5	4	6
Chiru.	028	69	F	PG				0	FGS 5	75	1	6	2	6
	034	59	M	FCG				0	FGS 5	75	2	6	2	6
	041	75	M	PD		Ant.		0	FGS 5	75	2	6	2	8
Ch. EP	022	95	F	FCD				0	THI 5	67	1	5	1	6
	043	39	F	FCG				0	THI 5	75	2	6	2	6
	052	60	F	FCD				0	THI 5	75	1	2	1	3
Immob.	032	50	M	PG	Voyage avion		OUI	0	FGS 5	75	2	3	2	3
	039	70	F	FSG	Voyage avion			0	FGS 5	69	2	6	2	6
Im. EP	046	86	M	FSD	Allitement			0	THI 5	70	1	3	1	4
	049	85	F	PG	Allitement			0	THI 5	75	1	2	1	2
	053	60	M	FCD	Allitement			0	THI 5	70	2	4	2	4
	056	16	F	FCG	Allitement			0	THI 5	75	0	0	4	7
								1	THI 5	66	2	4	2	4
Gé. EP	026	77	M	FSG		Ant.		0	FGS 5	75	2	6	2	6
	024	73	M	FCG		Ant.	OUI	0	FGS 5	73	1	6	1	6
	029	67	M	FSD		Ant.	OUI	0	FGS 5	75	1	1	2	2
Id. EP	027	78	F	PD				0	FGS 5	75	3	2	2	2
	031	87	M	FSG				0	FGS 5	75	2	2	2	2
	044	89	F	FCD				0	THI 5	72	2	7	2	7
	054	48	M	PG			OUI	0	THI 5	69	1	2	1	2
	040	54	M	PG				0	FGS 5	73	1	7	1	7
Idio.	042	92	F	FCD				0	THI 5	75	2	6	2	7
	047	30	M	FCG				0	THI 5	75	2	5	2	6
	045	39	M	FCD				0	THI 5	75	0	3	1	3
								1	THI 5	75	1	3	1	3
								3	THI 5	75	1	2	1	2
	048	62	M	FCD			OUI	0	THI 5	75	1	3	1	3
	050	68	M	PG				0	THI 5	75	2	3	2	3
								1	THI 5	75	1	2	1	2

Tableau 1.5 – Liste des acquisitions par patient obtenu avec le protocole 3 : classes, identifiant (ID), âge, sexe (féminin/masculin), TVP (Commune, Droite, Fémorale, Gauche, Supérieure et/ou Poplitée), facteur déclencheur, prédispositions génétiques (antécédents Familiers, antécédents personnel ou hormonothérapie), tabac (Tab.), mois d'examen (M), mode et fréquence de l'échographie (Echo.), gain, nombre d'échographies et d'élastographies transversales (T) et longitudinales (L).

1.3.3 Acquisition des images de glandes salivaires

Le service de rhumatologie du CHRU de Brest utilise également le système Toshiba Aplio 500 et a notamment collecté des images échographiques mode B et élastographiques de glandes salivaires dans le cadre de la détection du syndrome de Gougerot-Sjögren¹.

Comme leur nom l'indique, les glandes salivaires secrètent la salive qui est indispensable à la parole, à la mastication, au goût, à la déglutition et à la bonne hygiène de bouche [20]. Il existe trois paires principales de glandes salivaires illustrées en figure 1.16. Les glandes parotides, situées en dessous des oreilles, sont les plus volumineuses. Plus petites, les glandes submandibulaires (ou sous-maxillaires) et sublinguales se trouvent respectivement sous la mâchoire et sous la langue. Pour produire plus d'un litre de salive par jour, ces trois paires principales sont accompagnées de glandes accessoires (microscopiques) disséminées dans la muqueuse buccale. Les glandes salivaires peuvent être le siège de différentes affections, de nature infectieuse, lithiasique (présence de calcul(s) dans le système canalaire excréteur des glandes salivaires), tumorale et/ou immunologique. Ces affections peuvent entraîner une augmentation, parfois douloureuse, du volume des glandes ou générer un trouble du fonctionnement, soit par insuffisance, soit par hyperfonctionnement.

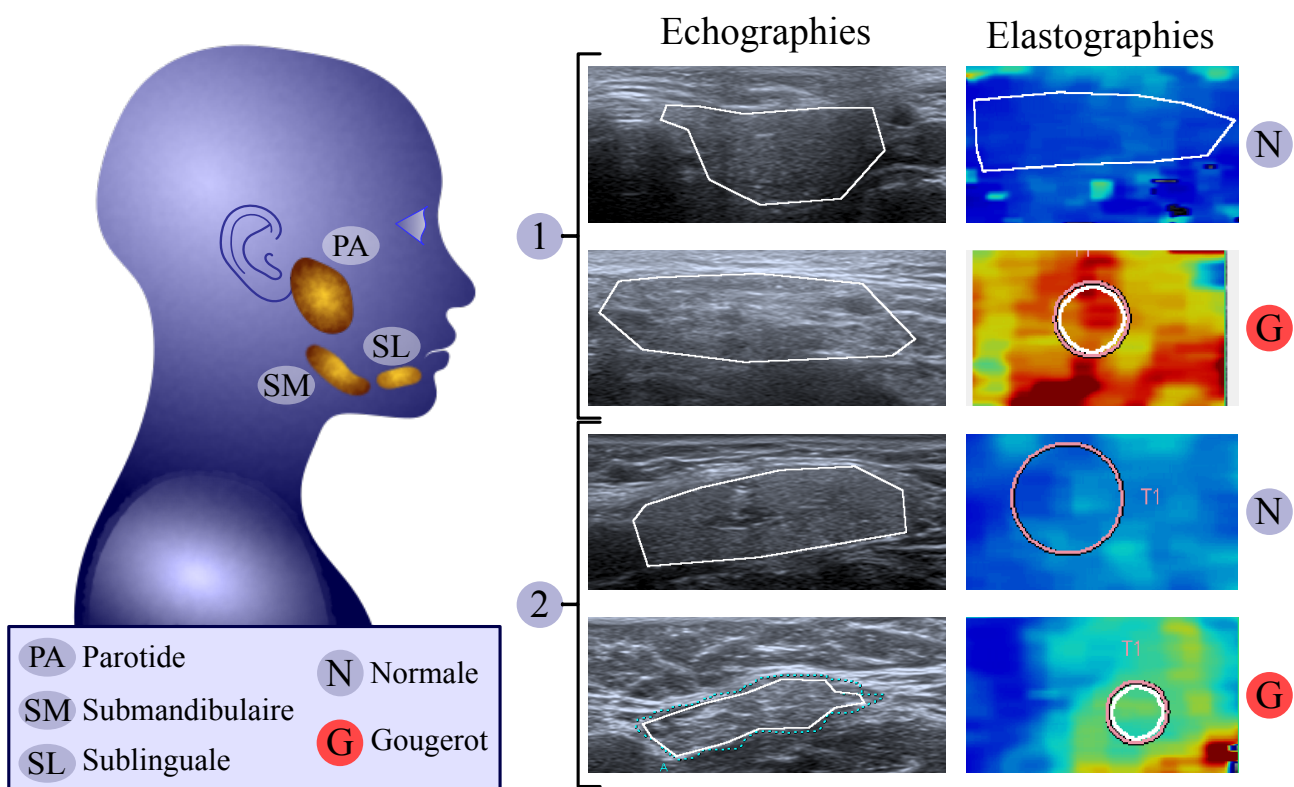


FIGURE 1.16 – Echographies et élastographies de glandes salivaires normales ou atteintes du syndrome de Gougerot.

Dans le cadre de notre projet, nous nous intéressons aussi, dans une moindre priorité, à la détection du syndrome de Gougerot-Sjögren [2, 47] qui est une maladie auto-immune : des cellules du système immunitaire, appelées lymphocytes, et des anticorps, sécrétés par ces mêmes cellules (lymphocytes), attaquent les glandes et provoquent leur dysfonctionnement. Cette maladie se traduit principalement par une sécheresse buccale et/ou oculaire (appelé syndrome sec) ce qui signifie que les glandes ne sont plus capables de produire suffisamment de larme et de salive [47]. Les patients examinés dans le cadre de cette étude sont généralement atteints d'un syndrome sec (fréquent à partir de 50 ans, notamment

1. Le syndrome de Gougerot-Sjögren est une maladie auto-immune qui se caractérise par une diminution des sécrétions de larmes et de salive réalisant un syndrome sec [47].

chez les femmes [41]) et consultent pour une suspicion du syndrome de Gougerot-Sjögren. En l'absence de cette maladie auto-immune, les glandes examinées sont considérées comme normales et cela signifie que la cause du syndrome sec est ailleurs. Comme on peut l'observer sur la figure 1.16, l'échographie d'une glande normale montre un tissu plutôt homogène alors qu'une glande atteinte du symptôme de Gougerot-Sjögren présente une texture beaucoup plus hétérogène et plus échogène. Comme pour les thromboses veineuses, des études [4, 80, 140] cherchent à montrer l'intérêt de l'élastographie pour le diagnostic de cette pathologie. En effet, les glandes malades, notamment les parotides, semblent être plus dures que les glandes normales comme l'illustre la figure 1.16.

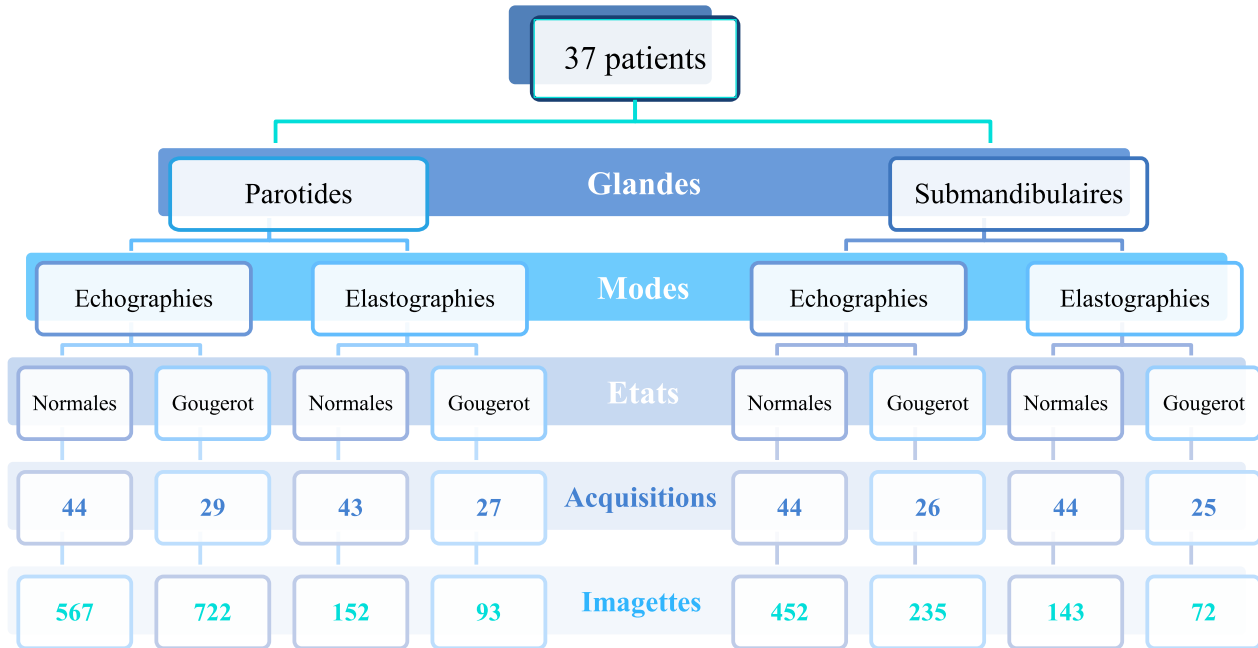


FIGURE 1.17 – Constitution de la base de glandes salivaires : nombre d'acquisitions et d'images extraites de ces acquisitions.

La figure 1.17 donne la composition de la base de glandes salivaires qui est utilisée dans les simulations présentées ultérieurement. Notre base est donc constituée de deux septantaines de glandes parotides et submandibulaires collectées sur 37 patients. Notre objectif est donc multiple :

1. La détection automatique du syndrome de Gougerot-Sjögren ;
2. L'intérêt de l'élastographie dans le diagnostic de cette maladie ;
3. L'évaluation de nos approches sur des images de même nature que celle des thrombi.

Néanmoins, des performances satisfaisantes sur les glandes salivaires (tissus) ne garantissent pas des résultats similaires sur les thrombi (structure gélatineuse). En effet, s'il est établi que les ultrasons se propagent bien dans les tissus humains, leur propagation dans un milieu mi-fluide, mi-solide est plus complexe à modéliser, notamment en ce qui concerne l'élastographie.

1.3.4 Conclusion sur les bases de données

Dans cette section, nous nous sommes intéressés à la problématique technique du sujet d'étude, à savoir comment collecter des données physiques *in vivo* (sur des patients) avec un système d'acquisition commercial performant et en constante évolution. Le système Toshiba Aplio 500 permet en effet d'obtenir des images haute-résolution de l'échogénicité et de la dureté du thrombus. Toutefois, ce système étant commercialisé et donc protégé, nous n'avons pas accès aux données brutes et n'avons qu'une connaissance limitée de l'ensemble des prétraitements effectués par le système pour générer l'image finale. C'est pour cette raison que nous avons tenté de standardiser au maximum nos acquisitions afin de limiter le biais pouvant être induit par un changement de paramétrage. En outre, des images de glandes salivaires collectées avec le même système nous permettront de valider nos approches avant de les appliquer aux images de thrombi veineux.

1.4 Conclusion partielle sur la partie médicale de notre projet

La difficulté du diagnostic de la thrombose veineuse profonde réside dans le manque de sensibilité (absence fréquente de symptôme) et de spécificité (communs à d'autres pathologies) de ses signes cliniques (*cf.* section 1.1). L'objectif de ce projet est de trouver des corrélations entre les facteurs déclencheurs (grossesse, immobilisation, cancer, *etc.*) et/ou la survenue d'une embolie pulmonaire avec la structure du thrombus. Pour analyser sa structure, deux techniques d'imagerie acoustique (présentées en section 1.2) sont envisagées : l'échographie et l'élastographie (carte de dureté). Ces deux techniques utilisent les propriétés des ultrasons et leur propagation dans le corps humain pour générer des images. En échographie, des ondes de compression hautes fréquences sont émises et l'image est formée à partir des échos réfléchis par les interfaces du milieu exploré. Quant à l'élastographie, elle consiste à étudier la propagation des ondes de cisaillement dans une zone d'intérêt car leur vitesse est directement liée à la dureté du milieu.

Notre base de données est donc constituée d'images échographiques et élastographiques qui ont été acquises avec le système Toshiba Aplio 500. La section 1.3 décrit la procédure de collecte des données et son évolution au cours du temps. En juin 2016, le système d'acquisition a été mis à jour dans le but d'améliorer la qualité des images. Dans le même but, un bras articulé est utilisé pour tenir la sonde contre la peau du patient et ainsi limiter la variabilité intra-opérateur. Pour caractériser la structure du thrombus, des acquisitions en coupes longitudinales ont été ajoutées au protocole initial qui considérait uniquement des coupes transversales du thrombus. Enfin, nous avons remarqué que la sonde convexe est plus adaptée que la sonde linéaire plus haute fréquence pour effectuer des élastographies de thrombus (mesures plus stables). Ces évolutions du protocole expérimental ont conduit à la création de trois bases de données. Avant de présenter les approches développées pour analyser les images collectées, le chapitre suivant fait le point sur les travaux existants sur l'imagerie du thrombus et plus généralement sur les méthodes de traitement d'images appliquées au biomédical.

Chapitre 2

Etat de l'art

Glossaire du chapitre 2

2D Deux Dimensions

3D Trois Dimensions

ACP Analyse en Composantes Principales (ou *Principal Component Analysis* en anglais, PCA)

NJW Algorithme de Ng, Jordan et Weiss

CS Classification Spectrale (*spectral clustering* en anglais)

DBSCAN Density-Based Spatial Clustering of Applications with Noise est un algorithme de partitionnement spatial à base de densité.

EH Egalisation d'Histogramme

EP Embolie Pulmonaire

IRM Imagerie par Résonance Magnétique

MAP Estimateur statistique du Maximum *A Posteriori* : méthode pour estimer des paramètres inconnus d'un modèle à partir de la vraisemblance et de la distribution *a priori* des paramètres.

Mode B Echographie en mode Brillance

MV Estimateur statistique du Maximum de Vraisemblance : méthode pour estimer des paramètres inconnus d'un modèle en construisant une fonction de vraisemblance (déduite de la fonction de densité) puis en maximisant son logarithme par rapport aux paramètres inconnus. Le MV est l'estimateur MAP pour une distribution *a priori* uniforme.

NGM Niveau de Gris Médian

OPTICS Ordering Points to Identify the Clustering Structure est un algorithme de partitionnement spatial à base de densité.

RN Réseau de Neurones

ROC Courbe permettant de mesurer la performance d'un classifieur (*Receiver Operating Characteristic* en anglais)

ROI Région d'intérêt (*Region Of Interest*)

SMA Classification par Sélection de Modèles Affines

SO *Scattering Operator* : réseau convolutif de transformée en ondelettes et d'opérateurs contractants

SVM Machine à vecteurs de support ou séparateur à vaste marge (*Support Vector Machine* en anglais)

SWIRE Shear Wave Induced Resonance Elastography

TCD Transformée en Cosinus Discrète

TEP Tomographie par Emission de Positons

TFD Transformée de Fourier Discrète

t-SNE t-distributed Stochastic Neighbor Embedding

TVP Thrombose Veineuse Profonde

Introduction

Lors du chapitre précédent, nous avons vu que la problématique du projet consiste à caractériser les thrombi veineux afin d'identifier la cause principale de leur formation (*i.e.* un ou plusieurs facteurs de risque) et d'estimer le risque d'évolution vers une embolie pulmonaire. Pour ce faire, nous disposons de deux technologies d'imagerie basée sur la propagation acoustique, l'échographie mode B et l'élastographie. La section 2.1 propose de faire un état des lieux des travaux existants sur la caractérisation du thrombus. Cette problématique étant encore peu développée et l'élastographie étant une technologie assez récente, la section 2.2 propose d'élargir le champ de l'état de l'art à d'autres applications biomédicales utilisant l'imagerie acoustique (voire Imagerie par Résonance Magnétique (IRM)) pour caractériser et classifier des tissus humains. Enfin, les sections 2.3 et 2.4 présentent des outils plus généraux et largement appliqués pour toutes problématiques d'extraction de caractéristiques et de classification.

2.1 Analyse du thrombus veineux

Cette section référence les publications de travaux sur la caractérisation du thrombus veineux. Si l'examen de référence en matière de Thrombose Veineuse Profonde (TVP) est la phlébographie, cet examen est rarement pratiqué (utilisation d'un produit de contraste, exposition aux rayons X, coût) et est remplacé par l'échographie (test de compression). Actuellement, afin d'étudier les propriétés viscoélastiques d'un thrombus, des images d'élastographie sont souvent effectuées en parallèle de l'échographie. Ce mode d'imagerie s'avère de plus en plus présent sur les systèmes d'échographie car il s'avère pertinent dans le diagnostic de pathologies du foie notamment (estimation de la dureté du foie). Son intérêt dans le diagnostic de TVP reste néanmoins à démontrer. Que ce soit par échographie ou élastographie, les travaux présentés dans la suite de cette section visent à dater le thrombus. En principe [150], le thrombus se développe en 3 phases :

1. Lors des 3 ou 5 premiers jours de son développement, le thrombus est encore mou et pas très bien fixé sur les parois de la veine. C'est pourquoi le risque d'embolie pulmonaire est plus élevé durant cette phase car un morceau du thrombus peut se détacher de la paroi et du reste du thrombus ;
2. Des jours 3 ou 5 au jour 14, le processus d'organisation du thrombus (qui devient de plus en plus fibrineux) est bien avancé le rendant plus solide et mieux fixé aux parois ;
3. A partir du jour 14, le thrombus est totalement organisé et il s'est légèrement rétracté par rapport à ses premiers jours d'existence.

Plus le thrombus est détecté tôt (dans les 7-10 premiers jours [155]), plus le traitement de la TVP va s'avérer efficace et plus le risque de séquelles post-thrombotiques sera faible. L'âge du thrombus joue donc un rôle crucial dans le choix (anticoagulants, thrombolytiques ou thrombectomie) et le succès du traitement de la TVP.

2.1.1 Par échographie

Cette sous-section se concentre sur les travaux qui cherchent à analyser l'échogénicité et l'échostructure du thrombus afin de le dater. Dans son mémoire de DESC (Diplôme d'Etude Spécialisé Complémentaire) de médecine vasculaire [14], S. Beuil s'intéresse au Niveau de Gris Médian (NGM) du thrombus sur une base de données de 20 patients dont l'âge des thrombi varie entre 0, 1, 3 et 6 mois. Les résultats ont montré que le thrombus semble devenir plus échogène avec l'âge ce qui est cohérent avec l'hypothèse que le thrombus s'organise et se solidifie avec le temps. Toutefois, les valeurs du NGM sont dépendantes du système et des paramètres d'acquisition. Pour être comparable, les acquisitions doivent être effectuées dans les mêmes conditions ce qui peut s'avérer compliqué suivant l'anatomie du patient (œdème, profondeur du thrombus) et rendre le thrombus invisible à l'échographie (nécessité donc de faire varier des paramètres pour voir le thrombus).

L'article [44] présente six méthodes de caractérisation du thrombus, c'est-à-dire six types de descripteurs, en vue d'une classification par réseau de neurone (apprentissage supervisé) :

1. **Histogramme** : les images échographiques sont représentées par leur histogramme (256 valeurs).

2. **Fenêtre glissante** : une région d'intérêt (30 x 30 pixels) est extraite de chaque échographie puis est balayée par une moyenne glissante (3 x 3 pixels). Ensuite, un vecteur contenant 100 valeurs est construit à partir des valeurs moyennes normalisées pour chaque image.

3. **Moyenne et variance** : chaque image est représentée par sa moyenne et sa variance.

4. **Seuillage** : les vecteurs descripteurs sont composés d'un ensemble de 0 et de 1 résultants d'un seuillage des images échographiques.

5. **Ondelettes** : les descripteurs sont issus de la décomposition en ondelettes biorthogonales de toutes les images.

6. **Matrices de cooccurrences** : après le calcul des 16 matrices de cooccurrences (pour 4 directions et 4 distances différentes), chaque image est représentée par l'énergie, l'entropie, l'inertie et l'homogénéité de ces 16 matrices.

Les auteurs de [44] comparent les six méthodes en fonction du nombre de descripteurs, du nombre d'itérations et du coût minimum. Les auteurs de [45] mesurent la similarité floue (« fuzzy similarity ») entre les descripteurs obtenus par les méthodes 3, 5 et 6 décrites précédemment. L'objectif est, ici, d'indexer (apparié) des images tests à partir d'une base d'apprentissage composée d'images de thrombi d'âges différents. Ces deux articles [44, 45] décrivent différentes méthodes d'extraction de descripteurs et de classification mais ne présentent pas de résultats de classification, faute d'une base de données suffisante.

Enfin, les auteurs de [96] ont développé un prototype d'imagerie échographique en trois dimensions (3D). La position de la sonde lors de son déplacement le long de la jambe est estimée à l'aide d'un capteur magnétique. Cette étude cherche à évaluer l'âge du thrombus à partir des images deux dimensions (2D) et trois dimensions (3D). Pour ce faire, plusieurs médecins estiment l'âge du thrombus à partir des échographies et du dossier en donnant leur degré de certitude (pas sûr du tout, pas très sûr, pas sûr, sûr, très sûr, absolument sûr). Ensuite, une décision sur l'âge du thrombus est prise à partir de leurs estimations de l'âge (classe) et du degré de certitude (converti en probabilité) en combinant la logique floue et un classifieur bayésien.

Les auteurs de [96] proposent de calculer le volume du thrombus au moyen de leur prototype d'imagerie 3D et grâce aux connaissances *a priori* sur la forme du thrombus. Pour ce faire, les contours 2D du thrombus sont approchés interactivement par une ellipse. Ensuite, la triangulation de Delaunay est appliquée à un ensemble de contours segmentés 2D positionnés en 3D pour construire en 3D la surface du thrombus. Des résultats préliminaires sont présentés pour sept thromboses : trois en plastique, trois *in vitro* et une *in vivo*. Le taux d'erreur sur le volume est inférieur à 4,5 % sur les thrombus plastiques et à 3,5 % pour les thrombus *in vitro*.

2.1.2 Par élastographie

Précédemment, nous avons vu qu'il y a actuellement peu de travaux sur la datation des thrombus veineux par échographie et aucun concernant leur classification suivant les facteurs déclencheurs de la thrombose. Le thrombus durcissant avec l'âge, l'approche par élastographie (carte de dureté) semble plus intuitive et fait donc l'objet de plus nombreuses publications : [116, 138, 142, 149, 150]. L'élastographie est une technique d'imagerie relativement récente puisque les premiers travaux dans le domaine datent des années 1980 et un des premiers systèmes d'élastographie a fait l'objet d'une publication en 1991 [58]. Les premières applications de l'élastographie ont eu lieu dans la détection de cancer de la prostate et la caractérisation de tumeur mammaire. Le premier système commercialisé d'élastométrie (mesure de dureté mais pas d'imagerie), le Fibroscan, a été développé en 2003 et devenu l'outil de référence pour la détection de pathologies du foie, remplaçant le diagnostic par palpation.

En 2005, Rubin *et al.* [138] ont comparé l'élastographie de deux thrombi *in vivo* : un subaigu (25 jours) et un chronique (3 ans). La technique d'élastographie utilisée est une méthode quasi-statique : la déformation est faite par pression de la sonde contre la peau du patient et, durant la déformation, 200 clichés échographiques sont acquis afin de suivre la déformation au cours du temps et d'estimer la dureté (module de Young) du thrombus en mesurant son déplacement par une technique complexe de corrélation entre les images. La dureté du thrombus a ensuite été normalisée par rapport à celle des veines. Les résultats ont montré une différence de dureté et de structure entre les thrombi subaigu et chronique. Le thrombus subaigu s'est révélé plus hétérogène et de dureté de 3 à 4 fois inférieure à celle des parois de la veine. Quant au thrombus chronique, il s'est avéré plus homogène et 10 fois plus dure que les parois. Toutefois, cette étude reste limitée par le faible nombre de cas (deux thrombi) et par le fait que les signes cliniques auraient ici suffi pour déterminer l'âge des thrombi.

Pour mettre en évidence le lien entre l'âge du thrombus et sa dureté, d'autres groupes de recherche se sont penchés sur l'analyse de thrombi animaux par élastographie. Siebers *et al.* [149] ont mené des études sur des thrombi porcins : après avoir provoqué une TVP chez des porcs, ces derniers sont euthanasiés et autopsiés à respectivement 1, 3, 6, 9, 12 et 15 jours (3 à 4 spécimens par âge de thrombus) afin d'en extraire le thrombus. Les thrombi sont ensuite placés dans un milieu gélatiné puis des élastographies semi-quantitatives sont menées. L'analyse de ces images (normalisées par rapport à l'élastométrie de la gélatine) a montré une forte augmentation de la dureté entre les âges 6 et 12 jours. De même, Mfoumou *et al.* ont observé l'évolution de dureté du thrombus sur des lapins en utilisant une méthodologie analogue. Les résultats des élastographies semi-quantitatives ont également montré que la dureté du thrombus a augmenté considérablement avec l'âge.

De plus, dans une seconde publication [150], Siebers *et al.* exposent une méthodologie afin de dater de façon automatique les thrombi porcins en deux catégories : moins et plus de 6 jours. La région d'intérêt ROI (*Region of Interest*) est définie par segmentation manuelle du thrombus sur l'échographie en Mode B (Brillance, voir section 1.2.2). La caractérisation des thrombi est menée par extraction de paramètres de texture (rapport signal à bruit, moyenne, écart type, contraste, entropie, *etc.*) de la ROI à partir de l'analyse spatiale et fréquentielle des élastographies. Le choix des paramètres pertinent a été effectué par analyse des courbes ROC (*Receiver Operating Characteristic*, pour « caractéristique de fonctionnement du récepteur ») puis la distance euclidienne, celle de Mahalanobis et le Maximum de Vraisemblance (MV) ont été utilisés pour classifier les thrombi en deux groupes. En appliquant une procédure de validation croisée des paramètres, les meilleurs résultats de classification ont été obtenus avec le MV (90% avec le MV, 77% avec la distance euclidienne et 80% avec celle de Mahalanobis).

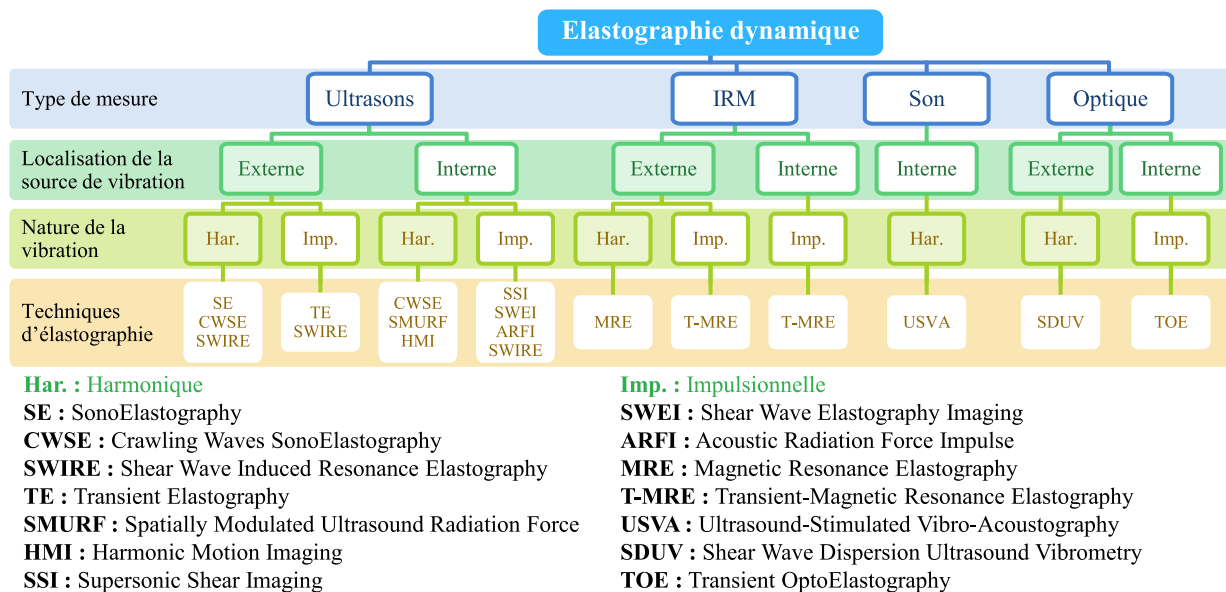


FIGURE 2.1 – Présentation simplifiée des techniques d'élastographie dynamique existantes [142].

L'imagerie par l'élastographie statique montre qu'il est possible de quantifier l'âge de thromboses. Néanmoins, malgré ses résultats prometteurs obtenus *in vitro* ou *in vivo* sur des thrombi animaux, cette technique présente des limitations importantes : elle est quantitative (déformation normalisée), opérateur-dépendante (contrainte externe qui diffère suivant l'utilisateur) et possède une sensibilité limitée (classement en thrombi récents ou très âgés). Pour pallier ces limitations, de nouvelles approches quantitatives, qui se basent sur l'élastographie dynamique, ont été proposées au cours des dernières années. Schmitt, dans son mémoire de thèse [142], fait un état de l'art des techniques d'élastographie dynamique existantes (*cf.* figure 2.1) et présente sa propre méthode nommée *Shear Wave Induced Resonance Elastography* (SWIRE). Les différences majeures entre les systèmes se résument principalement dans la façon dont les ondes de cisaillement sont générées (vibration, force de radiation, résonance, *etc.*) et pistées (échographie, optique, IRM, *etc.*).

Dans sa thèse [142], Schmitt compare aussi deux techniques pour estimer l'élasticité des thrombi, SWIRE et RheoSpectris. Ce dernier est un instrument de spectroscopie hyper-fréquence de la viscoélasticité de biomatériaux. Il permet de caractériser des matériaux dans une large gamme fréquentielle (entre 10 et 1000 Hz) par une analyse du spectre de vibration de l'échantillon soumis à une excitation mécanique impulsionnelle. Que ce soit avec des thrombi *in vitro* ou porcins *ex vivo*, les deux techniques ont donné des mesures d'élasticité équivalentes. En revanche, si l'utilisation de l'élastographie dynamique (pour dater les thromboses semble prometteuse par rapport aux techniques quasi-statiques (moins opérateur dépendant, mesures quantitatives), des résultats de caractérisation de la structure des thrombi *in vitro* ou *in vivo* ne sont actuellement pas présents dans la littérature. C'est pourquoi, cet état de l'art a été élargi à d'autres applications utilisant l'échographie et/ou l'élastographie dans le but de caractériser un tissu humain.

2.1.3 Conclusion sur l'analyse des thrombi veineux

Cette première section a fait un bilan des études menées en imagerie du thrombus. Nous pouvons constater que la problématique de datation du thrombus commence à s'enrichir avec notamment des expérimentations sur des animaux *in vitro*. Elle bénéficie également de l'arrivée de l'élastographie qui permet d'estimer la dureté du thrombus et d'en déduire son âge. Néanmoins, ces expérimentations sont difficilement applicables sur l'être humain (ligature de la veine cave pour l'observation *in vitro*) et l'une des limites actuelles de notre projet est qu'il est compliqué d'établir l'âge exact du thrombus (sauf en cas de facteur déclencheur évident comme une opération, un vol, une chute) et d'effectuer un suivi régulier de l'évolution du thrombus. Il est en effet souvent difficile de réexaminer le patient à intervalle régulier (une semaine par exemple) et, de plus, la progression de la thrombose est normalement freinée par le traitement anticoagulant ce qui peut induire un biais dans nos mesures. En revanche, la caractérisation des thrombi en fonction des facteurs déclencheurs et/ou la survenue d'une embolie pulmonaire reste encore peu étudiée donnant toute l'originalité de notre projet. C'est pourquoi, dans la section suivante, nous allons balayer d'autres applications médicales, où l'imagerie médicale est employée à des fins de caractérisation de tissus ou d'organes humains.

2.2 Méthodes de traitement d'images dans le domaine du biomédical

Dans un premier temps, nous allons chercher à faire le lien entre le type d'imagerie choisie et les domaines d'application dans la recherche médicale puis nous allons nous intéresser aux grandes familles de méthodes utilisées pour réaliser l'ensemble des étapes de la chaîne de traitement d'images (illustrée en figure 2.2) : prétraiter, segmenter, caractériser et classifier des images acoustiques (échographie ou élastographie).

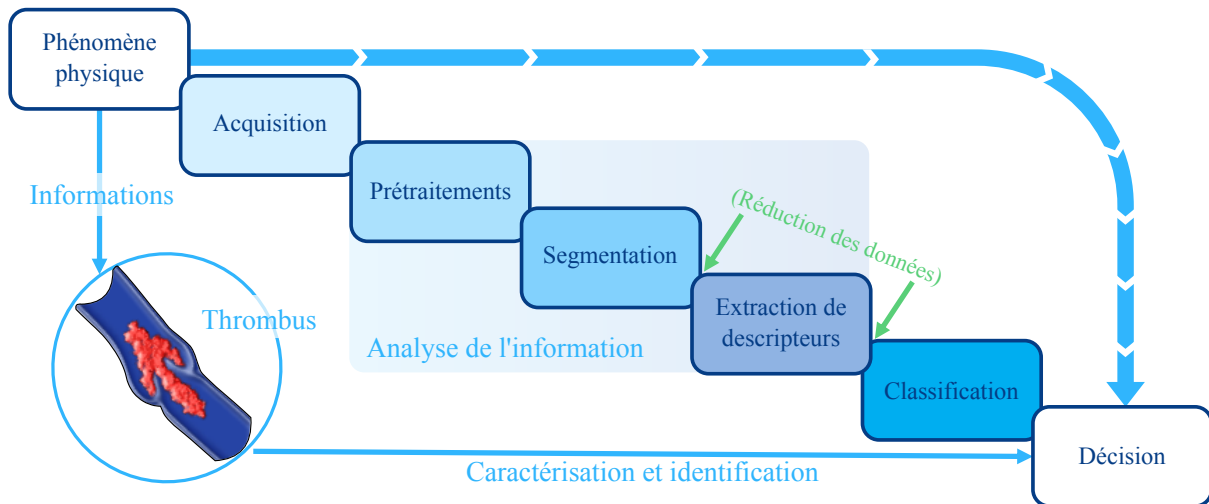


FIGURE 2.2 – Chaîne de traitement d'images : du phénomène physique à la prise de décision.

2.2.1 Domaines d'application

De nos jours, il existe quatre grands types d'imageries médicales [76] qui reposent sur l'utilisation des rayons X, des ultrasons, du champ magnétique ou de la radioactivité. La radiographie est principalement utilisée pour étudier les traumatismes osseux, les déformations du squelette ou les implantations dentaires. Elle est également pratiquée pour détecter des anomalies sur certains organes comme des infections bactériennes ou virales ou encore des tumeurs au niveau des poumons ou des seins (mammographie). Le scanner, à base de rayon X comme la radiographie, permet d'obtenir des images tridimensionnelles des organes ou des tissus et de suivre une modification de volume ou une anomalie de structure (infections, hémorragies, tumeur, ganglions, embolie, *etc.*). L'échographie, qui a l'avantage d'être non-invasive, non-irradiante et peu coûteuse, a pour principal application l'observation du fœtus, du cœur, du système vasculaire, des organes digestifs, urinaires et génitaux. Quant à l'IRM, c'est un examen coûteux, mais justifié lorsqu'un doute persiste après des radiographies, une échographie ou un scanner. Elle est principalement utilisée pour visualiser des « tissus mous » tels que le cerveau, la moelle épinière, les viscères, les muscles ou les tendons. Enfin, la découverte de la radioactivité a conduit au développement de la médecine nucléaire avec la scintigraphie puis la Tomographie par Emission de Positons (TEP). Ces techniques sont surtout utilisées en cancérologie [76].

L'IRM et la TEP sont des techniques d'imagerie plus récentes offrant une bien meilleure qualité d'image (résolution, précision, tridimensionnalité) et font donc l'objet d'un nombre important de projets de recherche. Toutefois, leurs disponibilités et leurs coûts font que ces outils ne sont utilisés qu'en cas de réel intérêt médical. En imagerie vasculaire par exemple, l'échographie (et le module doppler) suffit pour diagnostiquer et suivre l'évolution d'une thrombose veineuse. Par conséquent, il est actuellement très compliqué de justifier des IRMs de thrombus veineux, bien que cette technique permette d'observer la structure du thrombus de façon plus précise. C'est pourquoi cet état de l'art est principalement consacré aux techniques de détection, de caractérisation et de classification de tissus humains par ultrasons (échographie et élastographie). L'élastographie étant également une technique

récente, elle fait l'objet de nombreuses études notamment dans le diagnostic de lésions mammaires [153], de la prostate [42], du foie [27], de la thyroïde [119] et des glandes salivaires [80, 140]. La suite de cette partie énumère les différentes étapes de la chaîne de traitement (illustrée en figure 2.2) proposées dans la littérature pour prétraiter, extraire des descripteurs et classifier des images échographiques ou élastographiques. Elle se base notamment sur le papier de Cheng *et al.* [31] qui a réalisé une revue assez complète sur les techniques existantes pour classer des lésions mammaires. Dans la majorité des publications présentées, l'objectif principal est de classer les lésions en bénin/malin afin de limiter le nombre de biopsies.

2.2.2 Prétraitements

Les échographies sont principalement limitées par leur faible contraste et par la présence de bruit (*speckle*). Le but des méthodes du prétraitement est d'améliorer le contraste des images, notamment dans les régions d'intérêts (lésions) et de réduire le bruit sans perdre d'information utile pour la caractérisation du tissu ciblé.

Approches	Description	Avantages	Désavantage
Filtrages	Application d'un (ou plusieurs) filtre(s) dans le but de « nettoyer » le signal en éliminant le plus de bruit possible (additif, gaussien, flou, poivre et sel)	Simple et rapide	Sensible à la taille et la forme de l'image ; Possible perte d'information dans le domaine spatial
Ondelettes	Transformation dans le domaine des ondelettes et suppression du bruit (possibilité de transformation inverse)	Représentation du signal simplifié dans le domaine des ondelettes	Choix du seuillage dans le domaine des ondelettes
<i>Image compounding</i>	Moyennage d'images acquises à des fréquences et sous des angles différents	Simple	Augmentation du temps et de la complexité d'acquisition et de reconstruction

Tableau 2.1 – Approches de débruitage d'images échographiques (« speckle reduction ») [31].

On peut distinguer quatre grands types d'approches pour réduire le speckle : par Egalisation d'Histogramme (EH), par filtrages [101], par ondelettes, par moyennage d'échographies acquises à des fréquences et angles d'acquisition différents. Dans le cas du système Toshiba, cette dernière technique nommée en anglais « image compounding » est implémentée et les images de notre base de données en bénéficient certainement déjà. En ce qui concerne le filtrage (*cf.* tableau 2.1), on distingue deux familles de filtres :

- **Filtres linéaires** : moyennage, moyennage adaptatif (*i.e.* moyennage dans des régions homogènes)
- **Filtres non linéaires** : filtrage d'ordre deux (*e.g.* filtrage médian adaptatif), *Maximum A Posteriori* (*e.g.* MAP Pearlman Gauss), diffusion non linéaire (*e.g.* anisotropie), filtres géométriques, *etc.*

Quant aux ondelettes, la procédure de réduction de bruit se fait généralement en trois étapes : calcul de la transformée en ondelette, modification des coefficients en ondelettes pour supprimer le bruit et une transformée inverse pour obtenir l'image débruitée. Les techniques pour effectuer la deuxième étape peuvent être regroupées en trois groupes :

- « **Wavelet shrinkage** » : seuillage des coefficients (seuil fixe ou adapté à l'image) [28] ;
- **Approche bayésienne** : utilisation des statistiques des coefficients d'ondelettes [1] ;
- **Approche par filtrage ou diffusion** : utilisation des filtres décrits précédemment dans le domaine des ondelettes [134].

Les approches bayésiennes ou par filtrage surpassent les méthodes par seuillage dont la principale limitation reste le choix du seuil. Néanmoins, l'approche bayésienne suppose une connaissance *a priori* de la distribution du bruit, qui est en général inconnue en échographie. Si l'utilisation des ondelettes préservent en principe mieux les détails de l'image que les techniques de filtrage dans le domaine spatial, les passages du domaine spatial à celui des ondelettes et inversement augmentent la complexité en temps.

Afin d'améliorer la qualité d'une échographie, outre une réduction du bruit, il est possible d'améliorer le contraste de l'image afin de rendre certains détails ou contours plus perceptibles. La méthode la plus simple consiste à effectuer une EH [166] : des méthodes locales ou adaptatives permettent de ne pas perdre d'information dans les régions d'intérêt. En sous-section 3.2.3, ces méthodes d'EH, appliquées aux échographies de notre base de données, sont plus détaillées. De nombreuses publications suggèrent l'approche par la logique floue [67] qui, comme les méthodes par ondelettes, nécessite un changement de domaine de représentation (« fuzzification » et « defuzzification »).

2.2.3 Segmentation

Dans la chaîne de traitement d'images (visible en figure 2.2), l'étape suivant celle des prétraitements est la segmentation. Cette dernière est une étape charnière de la chaîne (l'une des plus difficiles en reconnaissance de forme mais essentielle) et elle consiste à délimiter une région d'intérêt (lésion ou tissu ciblé). Dans le cadre de notre projet, cette étape est effectuée à partir de l'ellipse dessinée autour du thrombus par l'opérateur qui fait l'acquisition. Toutefois, dans le cas où nous envisagerions de segmenter le thrombus uniquement à partir de l'échographie, le tableau 2.2 propose plusieurs types de techniques de segmentation tirées de la littérature.

Approches	Description	Avantages	Inconvénients
Seuillage d'histogramme [74, 79]	Après un lissage des images (le plus souvent), définition d'un ou plusieurs seuils qui permettent de séparer les tissus sains des lésions	Simple et rapide	Choix du seuil ; performances limitées dans de nombreuses applications
Modèles de contours actifs (<i>e.g.</i> Snake [82])	Après définition d'un contour grossier, déplacement du contour en fonction de la carte d'énergie jusqu'à épouser au mieux le contour de l'objet	Simple et indépendant de la forme de l'objet	Temps qui dépend du nombre d'itérations
Classification des pixels [34, 91, 97, 158]	Segmentation vue comme un problème de classification : classification spectrale, réseaux de neurones, <i>etc.</i>	Contours extraits de façon automatique	Définition du nombre de classes et/ou apprentissage
Champs de Markov [112, 125, 136]	Modélisation des textures ultrasonores par des champs markoviens conduisant à des lois de distributions locales (propres au tissu) représentant l'enveloppe du signal acoustique.	Précis et efficace	Couteux en calcul et en temps

Tableau 2.2 – Approches de segmentation d'images [31].

L'approche de segmentation la plus simple réside dans le seuillage de l'histogramme afin de conserver uniquement les niveaux de gris (ou les couleurs) caractéristiques de la zone à conserver [74, 79]. Comme cette approche est assez limitée, d'autres techniques s'appuyant sur une détection des contours (comme l'algorithme Snakes [82]) ou sur une classification sur les pixels de l'image [34, 91, 158] sont souvent privilégiées. Enfin, une quatrième approche, en général plus gourmande en temps de calcul et reposant sur les champs aléatoires de Markov, s'intéresse aux interactions locales entre niveaux de gris de pixels voisins [112, 136].

2.2.4 Extraction de descripteurs

L'objectif de cette étape est de décrire les données prétraitées et segmentées par des descripteurs (*features* en anglais). L'information extraite doit être la plus pertinente possible pour le type de classification souhaitée et la moins volumineuse possible pour limiter les ressources machines. En pratique, les données brutes (ou en sortie des prétraitements) des classes recherchées (origines des thrombi par exemple) sont plongées dans des surfaces très entremêlées. L'extraction de descripteurs joue un rôle clé puisqu'elle doit permettre de « détordre » ces surfaces (*manifold untangling* en anglais) afin que les classes soient plus facilement séparables (voir figure 2.3). Les transformations menant aux descripteurs sont irréversibles et peuvent être obtenues à partir de traitements simples (comme le sous-échantillonnage d'une image) ou plus complexes. Dans la littérature, il existe une large variété de descripteurs d'une image (cf. tableau 2.3) comme l'image elle-même, ses statistiques, son énergie, sa texture ou sa géométrie.

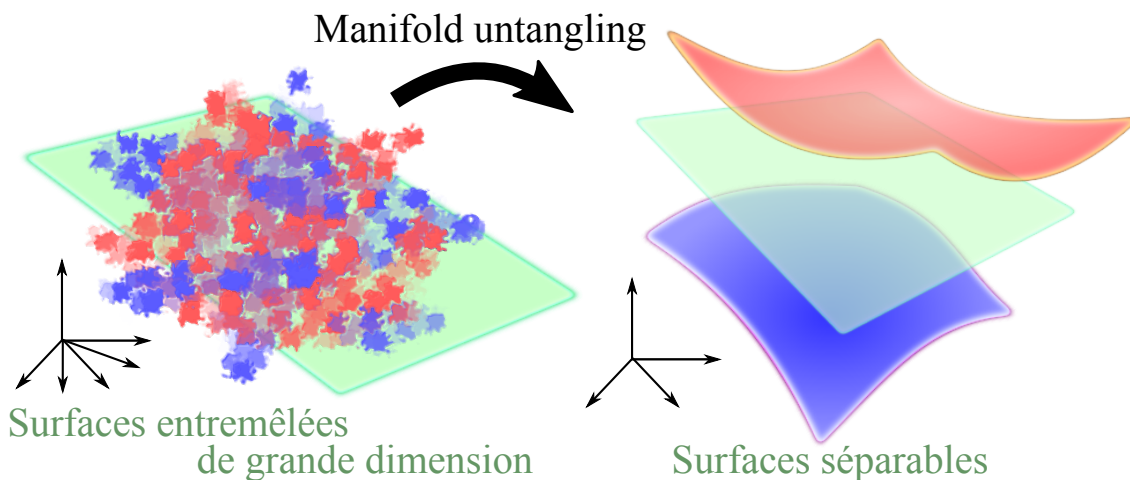


FIGURE 2.3 – *Manifold untangling* : l'extraction de descripteurs pertinents vise à détordre la surface des données afin de rendre les classes séparables (inspiré de [51]).

Approches	Exemples de méthode
Descripteurs statistiques :	Matrice de descripteurs statistiques (<i>Statistical Feature Matrix</i> SFM) [164]
	Coefficients d'autocovariance : description des corrélations entre pixels [75]
	Rapport entre la variance, les coefficients d'autocorrélation ou de la moyenne des tissus sains et ceux de la lésion [79]
	Matrice de cooccurrence [59, 118, 152]
	Statistiques du premier et du second ordre [81, 126, 152]
Descripteurs fréquentiels :	<i>Block Difference of Inverse Probabilities</i> (BDIP) [75]
	Transformée de Fourier Discrète (TFD) [152]
	Transformée en Cosinus Discrète (TCD)
	Fréquence relative d'éléments de contour [59]
Descripteurs de textures :	<i>Spiculation</i> : rapport entre les composantes basse-fréquences et haute-fréquences dans le domaine de Fourier [79]
	Energie, entropie, inertie et homogénéité locale [152]
	Analyse des coefficients en ondelettes : distribution, statistiques [117, 118]
Descripteurs morphologiques :	Contraste entre les niveaux de gris des pixels [79]
	Dissimilarités [79]
	Forme des contours de la lésion [30]
	Echogénéité des bords de la lésion [154]
Descripteurs basés sur des modèles :	Nombre de protubérances et de creux [30]
	Périmètre et aire de la lésion [30]
Descripteurs basés sur des modèles :	Distribution de type K, de Rayleigh ou de Nakagami [145, 146]
	<i>Power-Law Shot Noise model</i> (PLSN) [102]

Tableau 2.3 – Approches d'extraction de descripteurs [31].

Dans cette thèse, nous proposons par la suite deux approches pour décrire nos images de thrombi : le scattering operator (*cf.* section 4.1) et les statistiques d'ordre supérieur (*cf.* section 4.2). La première approche vise à caractériser la texture des images en utilisant les transformées en ondelettes et la seconde approche s'intéresse davantage à la distribution des pixels. La morphologie du thrombus, qui dépend de l'anatomie du patient, ne nous paraît pas être un critère approprié pour pouvoir identifier un ou plusieurs facteurs déclencheurs : l'analyse de la structure des thrombi nous paraît plus appropriée. En section 2.3, nous nous pencherons sur des outils qui permettent de réduire la taille des descripteurs tout en limitant la perte d'information.

2.2.5 Classification

A présent, l'objectif est d'analyser les descripteurs et de les trier en un certain nombre de classes (groupes d'images). Dans le cadre de notre projet, les classes souhaitées sont liées à l'épidémiologie des patients (grossesse, immobilisation, cancer, *etc.*) et la survenue d'une embolie pulmonaire. Le tableau 2.4 divulgue les principales approches de classification regroupées en deux catégories : les méthodes supervisées et non supervisées.

Classification supervisée (*i.e.* avec apprentissage) :

Approches	Exemples de classifieurs	Description
Classifieurs linéaires	Sélection de Modèles Affines (SMA) [22], Analyse Discriminante Linéaire (ADL) [146, 152]	Construction d'une représentation de faible dimension qui préserve les distances entre les données par combinaisons linéaires des variables originales.
Approche probabiliste	k plus proches voisins, classifieur de Bayes, réseau bayésien [81]	Classement à partir des probabilités <i>a priori</i> d'appartenance à une classe.
Arbres de décision	<i>Classification And Regression Trees</i> (CART) [100], <i>Inductive of Decision Trees</i> (IDT) [131]	Partitionnement hiérarchique selon des règles établies sur la base d'apprentissage.
Réseaux de neurones	<i>Artificial Neural Network</i> (ANN) [30, 154], <i>Deep Neural Network</i> (DNN) [35, 81, 97, 148]	Apprentissage sur un grand nombre de données pour estimer les meilleurs paramètres du réseau de neurones.
Approches statistiques	<i>Support Vector Machines</i> (SVM) [75, 133], Analyse en Composantes Indépendantes (ACI) [7, 39, 92]	Les données des classes sont représentées par quelques exemples (les points support) qui vont former des hyperplans utilisés ensuite pour séparer les données test.

Classification non supervisée :

Approches	Exemples de classifieurs	Description
Partitionnement par groupement de points	k -moyennes [72, 91], k -voisins [81], <i>Constrained k-means</i> [72]	Estimation du centre des groupes et rassemblement des données les plus proches de ces centres.
Partitionnement hiérarchique	<i>Waikato Environment for Knowledge Analysis</i> (WEKA) [70], <i>Robust Clustering using linKs</i> (ROCK) [66], <i>Cobweb</i> [55]	Décomposition hiérarchique (arbre) d'un ensemble de données basée sur des critères prédéfinis.
Partitionnement à base de densités	<i>Density-Based Spatial Clustering of Applications with Noise</i> (DBSCAN) [52], <i>Ordering Points to Identify the Clustering Structure</i> (OPTICS) [3]	Regroupement des points de proche en proche <i>i.e.</i> que deux points sont dans le même <i>cluster</i> si leur distance est inférieure à une certaine valeur.
Partitionnement à base de modèles	<i>Expectation Maximization</i> (EM) [91]	Classement à partir du modèle mathématique associé aux données.
Réseaux de neurones	Cartes auto-adaptives de Kohonen [84], Machine de Boltzmann restreinte, codage parcimonieux (<i>Sparse coding</i>)	Réseau auto-organisé : minimisation d'un certain critère, dictionnaire de motifs élémentaires.

Tableau 2.4 – Approches de classification des données.

La classification supervisée consiste à identifier la classe des nouvelles données (ou données tests) à partir de l'observation de données (ou données d'apprentissage). En pratique, la base de données est donc divisée en un ensemble d'apprentissage pour construire le modèle de classement et en un ensemble de test qui sert à évaluer le modèle de classement. Les principales familles de classifieurs supervisés sont référencées dans le tableau 2.4. L'idéal est donc de disposer d'une base de données suffisamment riche pour que la base d'apprentissage soit bien représentative du problème tout en conservant une quantité importante de données tests. Or, notre base de données de thrombi est actuellement de taille limitée et, par conséquent, nous nous sommes orientés dans un premier temps vers une classification non supervisée.

L'identification de groupes de données similaires dans un grand ensemble de données est devenue une problématique importante en analyse des données. De nos jours, la classification automatique (*data clustering*), qui consiste à regrouper les données les plus similaires entre elles (*cluster*), est appliquée dans de nombreux domaines comme la vision par ordinateur, le biomédical, la finance, les assurances ou la grande distribution. Dans la mesure où les notions de similarité et de groupe peuvent être explicitées de multiples façons, de nombreuses méthodes de classification automatique ont été proposées.

Le classifieur le plus utilisé, de par sa simplicité et son efficacité, est l'algorithme des k -moyennes (k -means). Néanmoins, pour obtenir de bonnes performances, cet algorithme nécessite que les sous-ensembles soient relativement distincts et de forme elliptique. Pour pallier cette contrainte sur la forme des sous-ensembles, d'autres méthodes de partitionnement ont été développées comme les méthodes hiérarchiques, à base de densité ou à base de modèle (voir tableau 2.4). Une des approches les plus répandues pour la classification non supervisée est le partitionnement spectral (ou *Spectral Clustering*) [71, 163] qui s'appuie sur l'utilisation du spectre (valeurs propres, vecteurs propres) de la matrice des similarités des données. Dans la littérature, de nombreuses techniques permettent d'obtenir une mesure de similarité entre deux objets. Le choix de cette mesure dépend essentiellement du domaine d'application et du type de données (variables numériques, binaires, textes). Nous reviendrons plus en détail sur certains algorithmes de classification en section 2.4.

2.2.6 Conclusion sur l'état de l'art en imagerie biomédicale

Précédemment, nous avons pu constater que la caractérisation des thrombi veineux est encore peu présente dans la littérature mais elle s'intensifie avec l'arrivée de l'élastographie qui paraît prometteuse, notamment pour la datation des thrombi. Dans cette section, nous nous sommes penchés sur d'autres applications mêlant échographie et élastographie. La détection et la caractérisation des lésions mammaires par échographie sont largement étudiées que ce soit par des outils statistiques, morphologiques et d'analyse de textures. L'élastographie, qui est devenue une technique de référence pour diagnostiquer une fibrose du foie, s'avère également efficace pour quantifier la dureté de la prostate. Dans cette section, nous avons également parcouru les grandes familles de méthodes pour prétraiter, segmenter, extraire des descripteurs et les classer. Nous reviendrons sur les prétraitements et la segmentation de nos images de thrombus au chapitre 3 et sur les approches développées pour les caractériser au chapitre 4. Dans les deux sections suivantes, nous allons nous focaliser sur les principaux outils pour réduire la taille des descripteurs et pour les classer automatiquement en fonction des facteurs déclencheurs de la thrombose et/ou de la présence d'embolie pulmonaire.

2.3 Réduction de la dimension des descripteurs

Sur la chaîne de traitement (*cf.* figure 2.2), les méthodes présentées dans cette section peuvent s'appliquer en amont ou en aval de l'extraction des descripteurs. Dans les deux cas, l'objectif est de réduire la dimension de l'espace tout en conservant l'information pertinente des données. La difficulté de la tâche est d'identifier cette information pertinente et de la dissocier de celle qui l'est moins et/ou qui est redondante. Nous pouvons choisir par exemple de sélectionner l'information de plus haute énergie dans le domaine de Fourier (*cf.* sous-section 2.3.1), ou de projeter les données dans un nouvel espace de représentation réduit en conservant l'information la plus énergétique (*cf.* sous-sections 2.3.2), ou encore de réduire l'espace tout en préservant les distances entre les images (*cf.* sous-section 2.3.3).

2.3.1 Réduction par passage dans le domaine de Fourier (ou des cosinus)

Une solution simple pour compresser un signal (ou une image) est de le réduire dans le domaine de Fourier puis d'effectuer la transformation inverse. La Transformée de Fourier Discrète (TFD) décompose le spectre du signal en fréquence élémentaires. Soit $s(t)$ un signal à temps continu supposé d'énergie finie :

$$\|s\|^2 = \int_{-\infty}^{+\infty} |s(t)|^2 dt < \infty \quad (2.1)$$

Le spectre (ou la transformée de Fourier) de $s(t)$ est défini par [128] :

$$S(f) = \int_{-\infty}^{+\infty} s(t)e^{-i2\pi ft} dt = \langle s(t), e^{i2\pi ft} \rangle \quad (2.2)$$

où $\langle a, b \rangle$ désigne le produit scalaire de a et de b . Dans le cas d'un signal discret $s(n)$ échantillonné en N points, sa TFD est donnée par :

$$S(k) = \sum_{n=0}^{N-1} x(n)e^{-i2\pi \frac{nk}{N}} \quad (2.3)$$

Un exemple de la transformée de Fourier d'un signal temporel est disponible en annexe B.1.

En règle générale, l'énergie du spectre est plutôt concentrée aux alentours de la composante continue (*i.e.* la fréquence nulle qui représente la valeur moyenne du signal) et est beaucoup plus faible vers les hautes fréquences. Dans le cas particulier où le signal est centré, la composante continue est donc nulle mais l'énergie du spectre reste plus importante en basse fréquence (sauf la fréquence nulle) qu'en haute fréquence. Le gain de compression vient de la réduction de précision des coefficients hautes fréquences (modification du pas de quantification) ou de leur suppression partielle (ceux d'énergie faibles). Dans le cas d'une image, comme nous pouvons le voir en figure 2.4, l'énergie du spectre est plus importante au centre de celui-ci et il suffit de supprimer les bords du spectre pour réduire la taille de l'image. En ajoutant des zéros sur les bords de l'image (*zero padding* en anglais), nous pouvons agrandir l'image sans ajouter d'information dans l'image. Réaliser un *zero padding* dans le domaine fréquentiel discret revient à augmenter le pas dans le domaine spatial discret, ce qui se traduit par un zoom de l'image dans le domaine spatial.

Une variante de la TFD, la Transformée en Cosinus Discrète (TCD), est très utilisée en compression du son ou de l'image, car elle possède une excellente propriété de concentration de l'énergie : l'information contenue dans sa transformée est principalement portée dans les basses fréquences. Au lieu de projeter le signal $s(k)$ sur une exponentielle complexe qui crée des coefficients complexes (TFD), la TCD le projette sur des cosinus de fréquences élémentaires et génère donc des coefficients réels. Par un procédé similaire à la TFD, la diminution (respectivement l'augmentation) de la taille d'une image consiste à passer dans le domaine des cosinus et à supprimer des lignes et des colonnes (respectivement ajouter des zéros) en bas et à droite (*cf.* figure 2.4).

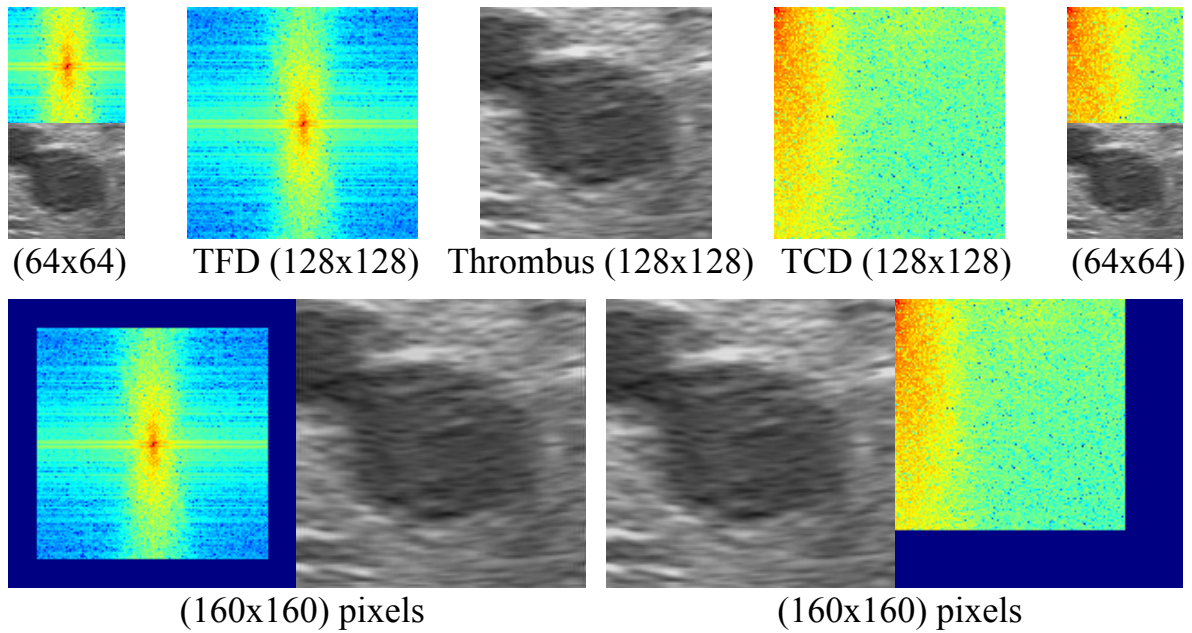


FIGURE 2.4 – Rétrécissement et agrandissement d'une échographie de thrombus de taille 128×128 par TFD et TCD.

2.3.2 Analyse en composante principale

L'objectif de l'Analyse en Composantes Principales (ACP) est de réduire l'espace de représentation des données (ou des descripteurs) en déformant le moins possible la réalité [78]. Elle consiste à transformer des variables corrélées (*i.e.* liées entre elles) en des nouvelles variables linéairement indépendantes (*i.e.* non corrélées). Ces nouvelles variables, appelées composantes principales (ou axes), peuvent être réduites en choisissant les axes les plus énergétiques au sens de la variance (*i.e.* de plus fortes inerties). Ces axes à fort pouvoir informatif permettent donc de générer un nouvel espace de dimension réduite. L'ACP est donc une approche à la fois géométrique (représentation des données dans un nouvel espace selon des directions d'inertie maximale) et statistique (recherche d'axes indépendants expliquant au mieux la variabilité des données).



FIGURE 2.5 – Projection de glandes salivaires sur deux axes par ACP.

Dans le cas d'une image, comme dans la figure 2.5, les pixels sont représentés dans un plan et peuvent être considérés comme une variable aléatoire à deux dimensions. En appliquant l'ACP sur l'image, on peut déterminer les deux axes expliquant le mieux la dispersion de l'objet (en l'occurrence une glande), qui est interprété comme un nuage de points. Le second axe est alors perpendiculaire au premier axe d'inertie maximale. Par cette approche, nous avons donc réduit notre image à $N \times N$ pixels à deux valeurs qui caractérisent la longueur et la largeur de la glande. Par ces paramètres, il serait sans doute possible de différencier les glandes parotides des glandes submandibulaires.

2.3.3 T-distributed stochastic neighbor embedding

En pratique, les données (brutes ou en sortie des prétraitements) sont de très grande dimension (nombre de pixels pour des images). Leurs descripteurs, bien que de dimension réduite, sont souvent supérieurs à trois ce qui rend la visualisation de la surface des données compliquée : comment projeter ces données de grandes dimensions vers des dimensions plus petites, pour lesquelles la visualisation sera possible et plus intuitive. Un choix possible est de sélectionner les deux ou trois premières composantes de l'ACP. Une autre méthode populaire, t-distributed Stochastic Neighbor Embedding (t-SNE), a été développée par Maaten et Hinton [104]. Elle propose de choisir un espace de basse dimension dans lequel la distance est préservée, en particulier pour les points proches dans l'espace de départ. Comme nous pouvons le voir dans l'exemple de la figure 2.6, la distance entre 2 points dans l'espace de départ sera similaire à celle dans l'espace réduit.

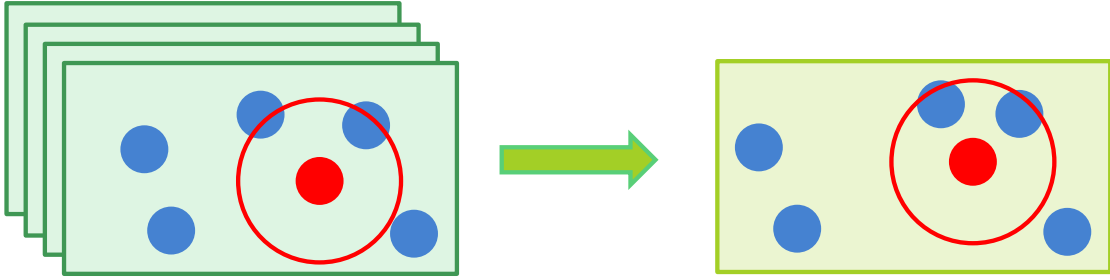


FIGURE 2.6 – Réduction de la dimension des données par la méthode t-SNE qui cherche à préserver la distance entre les données les plus proches.

D'un point de vue mathématique [104], la similarité p_{ij} entre deux données (x_i, x_j) dans l'espace de départ est modélisée par le rapport entre une fonction exponentielle et la distance quadratique entre les points dans l'espace d'entrée :

$$p_{ij} = \frac{\exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right)}{\sum_{k \neq l} \exp\left(-\frac{\|x_k - x_l\|^2}{2\sigma^2}\right)} \quad (2.4)$$

où $P = \{p_{ij}\}_{(i,j) \in D \times D}$ est la matrice de similarité entre les D données de la base, σ est le paramètre de dispersion de l'exponentielle et $\|x_k - x_l\|^2$ est la distance quadratique entre x_k et x_l calculée avec la norme $\| \cdot \|$ (e.g. la norme euclidienne).

Similairement, une distribution de Student (d'où le « t » de t-SNE) est utilisée pour représenter la similarité q_{ij} entre deux points (y_i, y_j) dans l'espace projeté (e.g. \mathbb{R}^2) :

$$q_{ij} = \frac{(1 + \|y_i - y_j\|^2)^{-1}}{\sum_{k \neq l} (1 + \|y_k - y_l\|^2)^{-1}} \quad (2.5)$$

où $Q = \{q_{ij}\}_{(i,j) \in D \times D}$ est la matrice de similarité entre les D données de la base.

L'espace projeté est construit en minimisant la divergence de Kullback Leibler $KL(P\|Q)$ (qui mesure la dissimilarité entre les distributions de probabilités). La fonction objectif C (loss function) est définie par :

$$C = \sum_i KL(P\|Q) = \sum_i \sum_j p_{ij} \log\left(\frac{p_{ij}}{q_{ij}}\right) \quad (2.6)$$

Le but de cette fonction C est de minimiser la distribution des distances entre les points dans les espaces de départ et d'arrivée, pour faire en sorte que, dans l'espace de faible dimension, les distances reflètent bien les distances dans l'espace initial de plus haute dimension.

En comparant la représentation des données avant et après l'étape d'extraction des descripteurs, l'algorithme t-SNE permet de visualiser le *manifold untangling* (cf. section 2.2.4) et de donner une indication sur la séparabilité des classes. Néanmoins, un nuage de points dans lequel nos classes ne seraient pas séparables ne signifie pas forcément que nos classes ne seront pas séparables avec un algorithme de classification puisque nous avons résumé de façon drastique nos descripteurs (réduction à deux ou trois dimensions). Enfin, en choisissant une dimension supérieure à trois, cet algorithme peut également être employé pour réduire la dimension des descripteurs.

2.3.4 Autres techniques

Il existe de nombreuses autres techniques de réduction de dimension [61] comme par exemple :

- **L'ACP à noyau** [143] : en utilisant l'astuce du noyau (en projetant les données dans un espace de plus grande dimension et en y effectuant des tâches de classification linéaire, les frontières de décision linéaires obtenues dans cet espace correspondent à des frontières non linéaires complexes dans l'espace d'origine), il est possible de réduire la dimension en réalisant des projections non linéaires complexes.

- **L'analyse discriminante linéaire** : proche de l'ACP, cet algorithme détermine les axes les plus discriminants pour définir un hyperplan sur lequel projeter les données.

- Le **plongement localement linéaire** (*locally linear embedding*, LLE) [137] : cette technique consiste à mesurer les relations linéaires entre chacune des données et ses plus proches voisins puis à trouver une représentation de plus faible dimension qui préserve ses relations locales.

- Le **positionnement multidimensionnel** (*MultiDimensionnal Scaling*, MDS) [87, 159] : dans le même principe que la méthode t-SNE, cette approche vise à réduire la dimension en préservant au mieux les distances entre les données.

- **Isomap** : contrairement aux méthodes t-SNE et MDS, celle-ci représente les données sous forme de graphe pour réduire la dimension tout en veillant au maintien des distances entre les nœuds.

- La compression dans le domaine des **ondelettes** : avec un procédé similaire à la TCD, elles permettent aussi de compresser les données et sont notamment utilisées pour la compression des images JPEG 2000.

2.3.5 Conclusion sur les techniques de réduction de dimension

Cette section a présenté quatre méthodes pour réduire la dimension des données tout en limitant la perte d'information : conserver l'information la plus énergétique dans les domaines de Fourier ou du Cosinus discrets, projeter dans un espace de dimension réduite en préservant les variables de plus hautes variances (ACP) ou les distances (t-SNE). En poussant la réduction à deux ou trois dimensions, ces deux dernières méthodes peuvent également servir d'indicateur de la séparabilité des classes de données. En effet, si les classes sont séparables dans l'espace des descripteurs projetés en deux ou trois dimensions, alors elles le seront sûrement aussi dans l'espace initial par un algorithme de classification. Bien sûr, d'autres approches existent pour réduire la dimension comme l'ACP à noyau, l'analyse discriminante linéaire, le plongement localement linéaire ou les ondelettes. Ces approches, plus complexes, n'ont pas été implémentées lors de la thèse car elles n'offrent pas forcément plus de garanties que des algorithmes plus simples (comme l'ACP ou la TCD). De plus, les méthodes de réduction de dimension sont appliquées, en règle générale, seulement lorsque nous obtenons des bonnes performances de classification sans réduction de dimension et que nous souhaitons améliorer l'efficacité des calculs (réduire le temps de calcul et son coût en mémoire). Dans la section suivante, nous nous pencherons sur des algorithmes de classification largement utilisés actuellement.

2.4 Classifieurs

Cette section se consacre à l'étape ultime de la chaîne de traitement à savoir la classification (*cf.* en figure 2.2). La section 2.4.1 présente deux approches supervisées classiques (une linéaire par Sélection de Modèles Affines (SMA) et une non linéaire avec les machines à vecteurs de support (*Support Vector Machine*, SVM) et l'approche en vogue de ces dix dernières années : l'apprentissage profond ou *deep learning* en anglais. De manière similaire, la section 2.4.2 décrit le principe de trois algorithmes populaires de *clustering* (*i.e.* non supervisés) : les k -moyennes, DBSCAN, et la classification spectrale.

2.4.1 Classification supervisée

La classification supervisée consiste à apprendre un modèle pour chacune des classes que l'on souhaite prédire et comparer les nouvelles données à chacun des modèles pour en déduire leur classe. Comme évoqué précédemment, cette section décrit trois méthodes populaires parmi tous les algorithmes de classification supervisée existants : une approche linéaire (SMA) et deux approches non linéaires (SVM et l'apprentissage profond).

2.4.1.1 Classification par Sélection de Modèles Affines (SMA)

L'Analyse en Composantes Principales (ACP), présentée précédemment pour réduire l'espace de représentation des descripteurs en déformant le moins possible la réalité (*cf.* section 2.3.2), peut également être utilisée pour générer des modèles affines de chaque classe. Le principe de la classification supervisée (par SMA) est illustré en figure 2.7. Pour commencer, il faut diviser la base de données en deux sous-ensembles : l'un pour l'apprentissage et l'autre pour le test. Lors de l'apprentissage, l'ACP (ou un autre algorithme comme l'analyse linéaire discriminante ou l'ACP à noyau) est utilisé pour sélectionner le modèle affine des différentes classes en résumant le plus pertinemment possible les descripteurs : chaque classe est modélisée par un centre, la moyenne de ses descripteurs, et par un certain nombre des nouvelles variables non corrélées à forte variance (déduts des descripteurs par l'ACP). Le choix du nombre d de ces nouvelles variables est à définir.

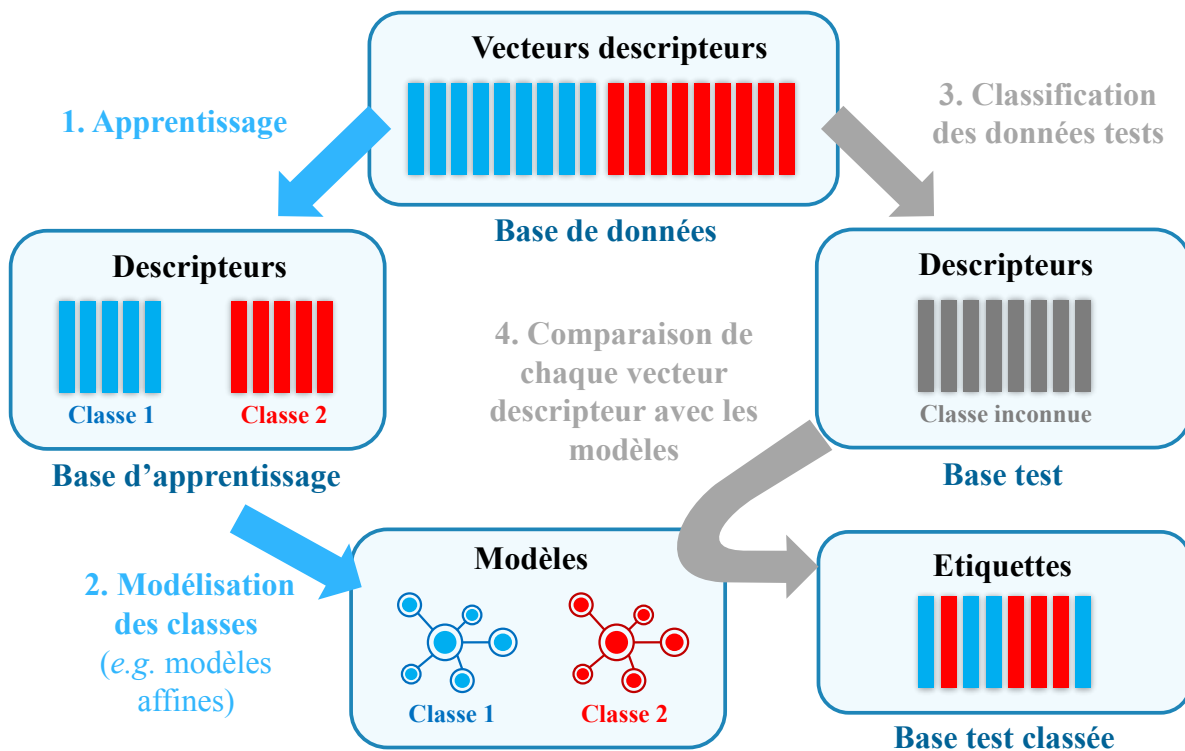


FIGURE 2.7 – Principe de la classification supervisée (par SMA par exemple).

Une fois l'ensemble des classes modélisées, on peut passer à l'étape de classement des descripteurs de la base de test. Chaque vecteur est projeté sur les espaces réduits de chacune des classes (*i.e.* les modèles affines) et est associé à la classe la plus proche (*i.e.* dont la distance entre le vecteur projeté et le centre du modèle est la plus faible). D'un point mathématique [22], si l'on note S_I le vecteur descripteur d'une donnée I de la base test, alors I sera associée à la classe k par la décision :

$$k(I) = \arg \min_{k \leq K} d_{euc}(S_I - P_{A_{d,k}}(S_I)) \quad (2.7)$$

où K est le nombre de classes, d_{euc} la distance (par exemple euclidienne), $P_{A_{d,k}}(S_I)$ est la projection de S_I dans l'espace de la classe k et paramétré par le nombre de dimensions d . Cette règle de décision correspond à la distance minimale entre le vecteur I et un espace affine décrivant les déformations et variabilités structurelles des descripteurs d'une même classe [22]. D'un point de vue pratique, le paramètre d (dimension de l'espace affine d'une classe), doit être optimisé.

2.4.1.2 Support Vector Machine

Les machines à vecteurs de support (*Support Vector Machine*, SVM) sont une généralisation des classifieurs linéaires comme la SMA : elles sont capables d'effectuer des classifications binaires non linéaires très puissantes [93]. Conçues dans les années 1990 [61], elles ont remplacé dans de nombreuses applications les systèmes de réseaux de neurone (que nous étudierons dans la section suivante) beaucoup plus coûteux en temps et en données d'apprentissage. Le principe fondamental des SVM consiste à construire un chemin de largeur maximale entre deux groupes d'observations (*i.e.* deux classes) et à utiliser ce chemin comme frontière de décision pour les données à classer. Leur particularité par rapport aux classifieurs linéaires est de remplacer un problème d'origine complexe par un problème simplifié au prix d'une augmentation de la dimension comme illustré en figure 2.8.

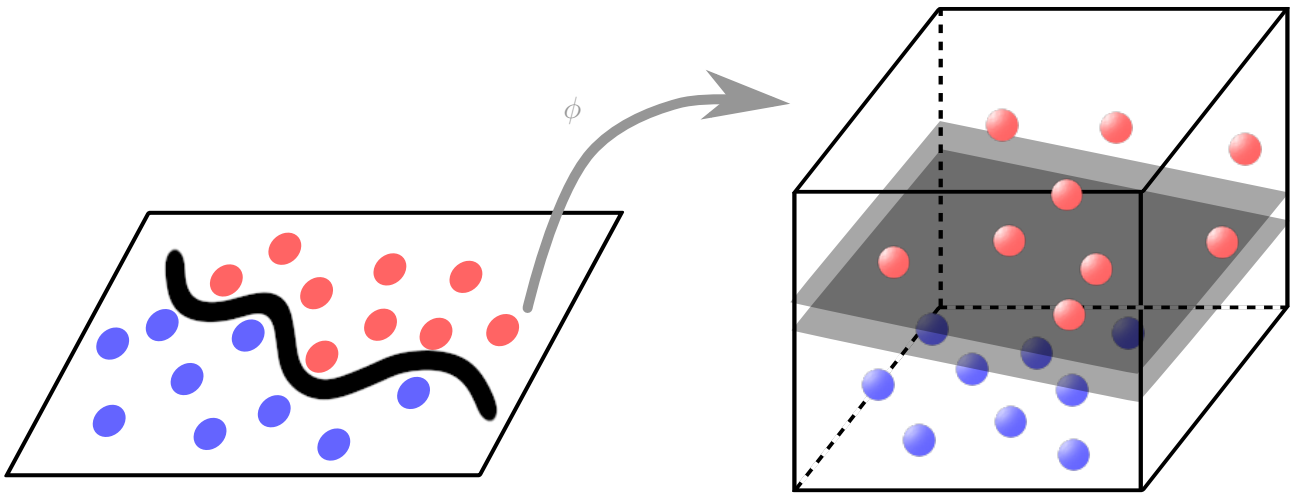


FIGURE 2.8 – Astuce du noyau SVM : projection des données dans un espace de plus grande dimension dans lequel elles seront linéairement séparables (inspiré de [93]).

Pour changer d'espace, l'idée générale est d'appliquer une transformation ϕ non linéaire qui envoie les données d'apprentissage $x_i \in \mathbb{R}^p$ de dimension p dans un nouvel espace $\phi(\mathbb{R}^p)$ de dimension supérieure à p dans lequel les classes seront plus facilement séparables. Si trouver une transformation ϕ appropriée peut s'avérer difficile, l'utilisation des SVM ne nécessite pas de connaître explicitement ϕ : l'astuce consiste à calculer une bande de séparation linéaire de largeur maximale en faisant intervenir uniquement les distances et les produits scalaires entre les points à séparer (leurs coordonnées explicites ne sont donc pas nécessaires). Si l'on note $\phi(x_i)$ les images des points x_i de l'espace d'origine, il suffit de connaître la fonction $K(x_i, x_j) = \langle \phi(x_i), \phi(x_j) \rangle$ qui donne les produits scalaires entre les points

transformés. Cette fonction clé K , appelée noyau, n'est pas sans rappeler l'astuce qui permet à l'ACP de classer des données non linéaires évoquée en section 2.3.4. Les noyaux les plus courants [61] sont :

$$\text{Linéaire : } K(x_i, x_j) = x_i^T \cdot x_j \quad (2.8)$$

$$\text{Polynomial : } K(x_i, x_j) = (\gamma x_i^T \cdot x_j + r)^d \quad (2.9)$$

$$\text{Radial gaussien : } K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2) \quad (2.10)$$

$$\text{Sigmoidé : } K(x_i, x_j) = \tanh(\gamma x_i^T \cdot x_j + r) \quad (2.11)$$

où l'exposant T indique la transposée et $\|x\|$ la norme de x (e.g. la norme euclidienne).

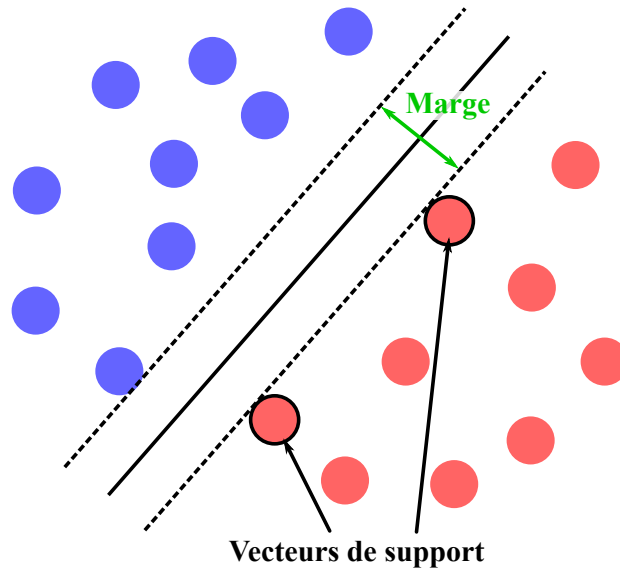


FIGURE 2.9 – Principe d'un SVM : création d'un chemin avec la plus grande marge possible entre les deux classes.

Dans un problème 2D linéaire (projection de l'espace transformé par exemple), le choix de la bande linéaire est illustrée sur la figure 2.9. Un classifieur SVM ajuste la bande la plus large possible (matérialisée par les deux lignes pointillées) d'où l'appellation française de « Séparateurs à vaste marge ». La marge est la distance entre l'hyperplan séparant les deux classes et les observations les plus proches de cet hyperplan. Ces observations sont appelées vecteurs de support d'où le nom de « machine à vecteurs supports ». Les SVM sont sensibles aux différences d'échelle des variables (palliable par une normalisation) et aux données aberrantes : les données dans l'espace transformé doivent être linéairement séparables. Pour éviter ce dernier problème, il est possible de relâcher la contrainte en autorisant un certain nombre de points à se trouver à l'intérieur de la bande, voire même du mauvais côté. On parle alors de classification à « marge souple » (*soft margin*) par opposition à « marge rigide » (*hard margin*).

2.4.1.3 Apprentissage profond

Depuis plusieurs années, les technologies qui reposent sur l'intelligence artificielle envahissent progressivement notre quotidien [93] : reconnaissance vocale (*Siri*, *Cortana* ou *Google now*), identification faciale, traduction automatique, jeux vidéos, robots autonomes, conduite automatique. L'idée des Réseaux de Neurones (RN) n'est cependant pas nouvelle puisque les premières tentatives datent des années 1960. Il faut attendre vingt ans et les premiers algorithmes de rétro-propagation de l'erreur pour voir les premiers succès des RN. Cependant, la puissance de calcul nécessaire pour entraîner ses réseaux et l'arrivée de méthodes alternatives comme les SVM ont contribué à la mise en sommeil des RN [65].

Le principe des RN consiste à s'inspirer du fonctionnement du cortex visuel des mammifères [73, 93]. Dans une version simple, chaque neurone d'un tel réseau possède un niveau d'activation compris entre 0 et 1. Le schéma d'interconnexion entre les neurones définit l'architecture du réseau. Une architecture classique consiste à organiser les neurones en couches successives avec des interconnexions limitées aux couches adjacentes comme le montre la figure 2.10. Un exemple classique est le classement automatique de chiffres manuscrits. Les niveaux d'activation de la couche d'entrée correspondent aux niveaux de gris des pixels (ou aux descripteurs) de l'image d'entrée. La couche de sortie indique la probabilité que cette image soit chacun des dix chiffres. L'entraînement du réseau consiste à trouver les poids des neurones des couches cachées telle que la couche de sortie classe avec précision les images de la base d'apprentissage.

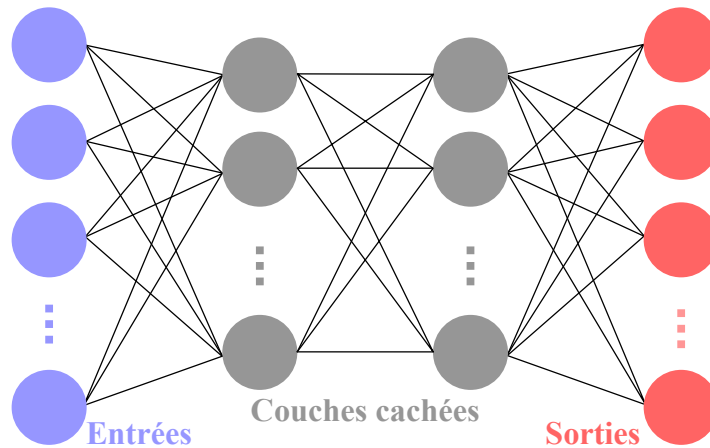


FIGURE 2.10 – RN organisé en quatre couches dont deux cachées (inspiré de [113]).

L'émergence de nouveaux modèles pour apprendre de manière non supervisée les réseaux profonds (*Deep Belief Nets*) combinée à la croissance exponentielle de la puissance de calcul des machines (*CPU, GPU*) a réveillé l'intérêt de la communauté scientifique [141]. Le premier résultat marquant est à mettre au crédit de *Microsoft* lors d'un challenge gagné pour de la reconnaissance vocale. Le second succès très important a eu lieu lors du challenge *ImageNet* (un million d'images étiquetées réparties en mille catégories) par le réseau *AlexNet* [86]. Dans de nombreux domaines, l'arrivée des RN convolutifs comme *AlexNet* a permis d'obtenir un saut de performance spectaculaire par rapport aux modèles de références.

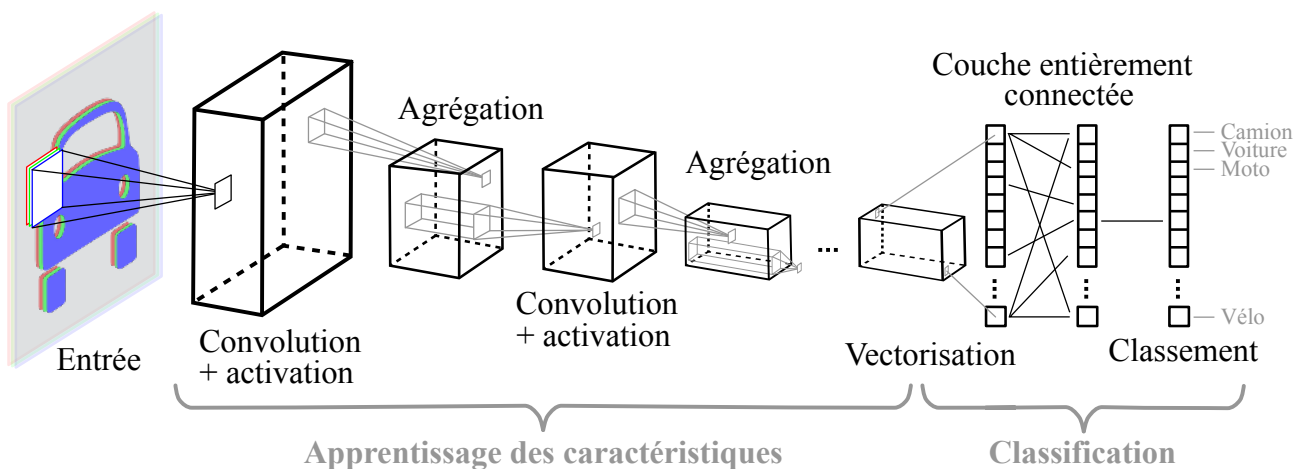


FIGURE 2.11 – Architecture du RN convolutif *ConvNet* de Matlab : chaque image de la base d'apprentissage passe dans les différentes couches du réseau qui sont entraînées de manière à décrire de façon précise chaque classe. Il suffit ensuite de passer une image de test dans le réseau pour obtenir en sortie la probabilité d'appartenance à une classe (créé à partir de [113]).

Depuis 2012, beaucoup de bibliothèques ont été développées par la communauté pour générer leur propre RN adapté à leur application [65]. Il y a, par exemple, *ConvNet* [162], écrite en *MATLAB* historiquement et mise à disposition par l'université d'Oxford (dont l'architecture est schématisée en figure 2.11). Il y a *Caffe*, *Caffe2* et *DeCAF* développées d'abord par *Berkeley Vision and Learning Center* (BVLC) et désormais maintenues par *Facebook*. Ce dernier avait d'abord proposé *Torch* qui a récemment vu sa version interfacée avec *Python* (*PyTorch*). On trouve également *Theano* (*Python*), développée par l'université de Montréal, et *TensorFlow* originellement conçue par Google. Une autre bibliothèque populaire est la bibliothèque *Keras*, créée par *Google*, qui est une surcouche des bibliothèques *TensorFlow* et *Theano*.

A présent, nous allons nous intéresser d'un peu plus près à un RN à convolution, le réseau *ConvNet* qui est schématisé en figure 2.11. Dans un tel réseau, il existe principalement quatre types de couches [63, 86, 113, 162] :

- La couche de **convolution** passe les données d'entrée par un ensemble de filtres convolutifs qui ont pour rôle d'extraire certaines caractéristiques des entrées (descripteurs). Les premières couches de filtres peuvent extraire des caractéristiques très simples, comme par exemple la luminosité ou les bords, alors que les couches profondes vont plutôt faire ressortir des caractéristiques plus complexes qui vont définir l'objet de façon unique. Contrairement aux méthodes traditionnelles, les filtres de convolutions ne sont pas prédéfinis selon un formalisme particulier (comme pour le scattering operator que nous verrons en section 4.1), mais ils sont appris par le réseau lors la phase d'entraînement : ils sont d'abord initialisés puis mis à jour par rétro-propagation du gradient. Les noyaux des filtres désignent les poids des couches de convolution.

- La couche d'**activation** suit généralement une couche de convolutions et correspond à l'application d'une fonction non linéaire aux descripteurs extraits à la couche précédente. C'est l'ajout de cette non linéarité qui va permettre de résoudre des problèmes de classification avec des frontières de décisions complexes. Une fonction d'activation populaire est la fonction **ReLU** (Rectified Linear Units) qui remplace toutes les valeurs négatives reçues en entrées par des zéros. L'ajout de ce type de couche permet également d'accélérer et d'optimiser l'apprentissage car seuls les neurones activés sont transférées dans la couche suivante.

- La couche d'**agrégation** (ou *pooling* en anglais) est une étape importante pour réduire la taille spatiale des cartes de descripteurs (sorties des couches de convolution-activation) par l'intermédiaire d'un pas, *stride* en anglais, qui permet d'ajouter une invariance partielle en translation et une stabilité accrue des représentations internes. Parmi les opérateurs usuels d'agrégation, il y a la fonction *max*, la moyenne ou la norme L_p (e.g. norme euclidienne si p égal à 2). Cette couche permet d'améliorer l'efficacité du réseau en réduisant le nombre de variables du réseau et de limiter les risques de sur-apprentissage.

- La couche **entièrement connectée** (*fully-connected*) suivie constitue bien souvent l'avant dernière couche d'un RN (convolutif ou non). Ce type de couche reçoit un vecteur en entrée d'où la couche de vectorisation (*flatten*) en amont. Par combinaisons linéaires, elle produit un vecteur de sortie qui indique la probabilité pour l'image d'entrée d'appartenir à chacune des classes. Pour calculer ces probabilités, la couche entièrement connectée multiplie chaque élément en entrée par un poids, fait la somme, puis applique une fonction d'activation (*softmax* par exemple). La matrice des poids est, comme les filtres de convolution, apprise lors de phase d'entraînement du réseau.

Un RN à convolution peut avoir plusieurs dizaines, voire centaines, de couches qui apprennent chacune à identifier différentes caractéristiques d'une image. Pour une application donnée, il y a généralement deux alternatives d'apprentissage des filtres du réseau : partir de zéro ou utiliser un modèle pré-entraîné. Dans le premier cas, l'architecte du réseau doit définir le nombre de couches et de filtres, ainsi que d'autres paramètres réglables. Il doit également disposer de grandes quantités de données, parfois des millions d'échantillons. L'autre alternative consiste à utiliser un modèle pré-entraîné pour extraire automatiquement des caractéristiques à partir d'un nouveau jeu de données. Cette méthode, appelée apprentissage par transfert, est une solution pratique pour appliquer

l'apprentissage profond sans avoir à recourir à un très grand jeu de données ni à procéder à des calculs et un apprentissage très longs. Notre base de thrombi étant relativement petite (quelques dizaines de patients), il faudrait probablement se tourner vers ce type d'alternative, qui a déjà donné de bonnes performances en imagerie médicale [23, 98, 148].

2.4.2 Classification non supervisée

La classification non supervisée cherche à regrouper les données les plus proches les unes des autres sans connaissances en amont. Pour que ce genre d'approches fonctionnent, il est souvent nécessaire de transformer l'espace de représentation des données (réduction de dimension et/ou extraction de descripteurs). Cette section présente le fonctionnement des trois approches utilisés pour classifier nos images de thrombi : deux algorithmes de partitionnement et une méthode qui transforme l'espace avant d'appliquer un algorithme de partitionnement.

2.4.2.1 Partitionnement par k -moyennes

Le k -means (ou k -moyennes) [72] est l'algorithme de partitionnement (*clustering* en anglais) le plus populaire de par sa simplicité et son efficacité. Son principe, illustré en figure 2.12, est le suivant :

1. Partitionnement aléatoire des points en k sous-ensembles (*clusters*) et calcul des centroïdes (*i.e.* centre) de chaque cluster ;
2. Tant que les centroïdes ne sont pas stabilisés (*i.e.* ne convergent pas vers un minimum local) :
 - (a) Calcul de la distance entre chaque point et chaque centroïde ;
 - (b) Attribution de chaque point au plus proche cluster (*i.e.* dont la distance avec son centre est la plus petite) ;
 - (c) Mise à jour des centroïdes de chaque cluster.

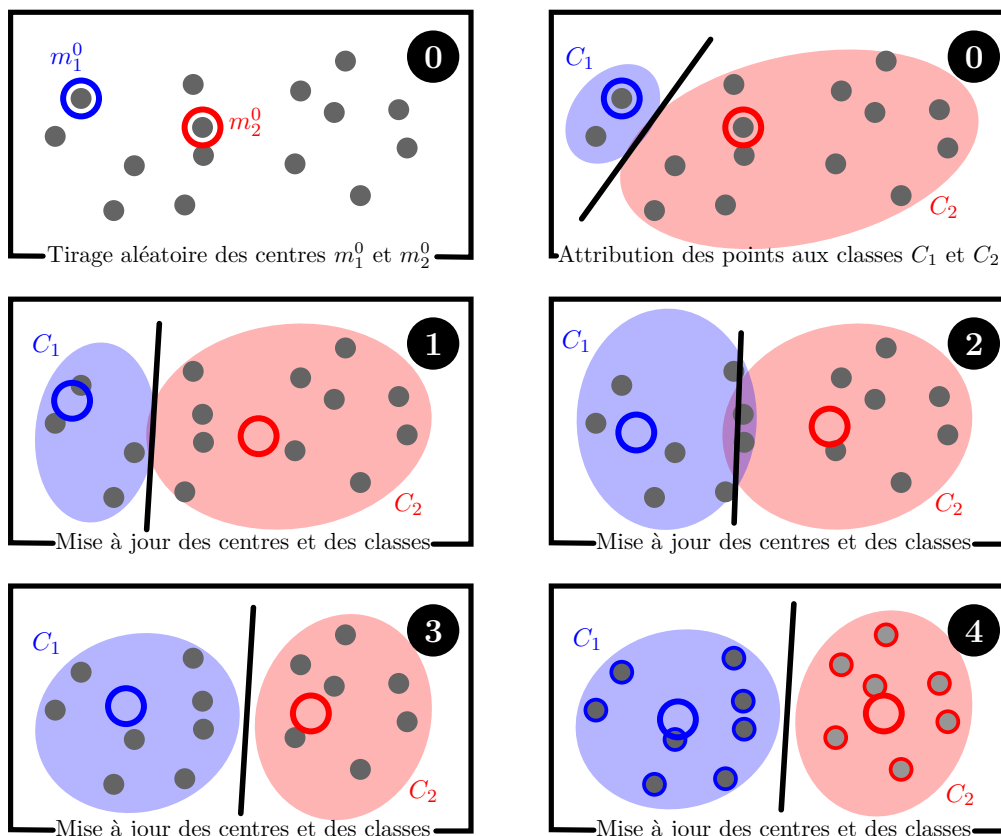


FIGURE 2.12 – Itérations successives de l'algorithme des K -moyennes en deux classes.

En pratique, l'algorithme s'arrête lorsque deux itérations successives conduisent à la même partition ou lorsqu'un certain critère est atteint. Les deux critères les plus courants sont le nombre maximal d'itérations et le nombre minimal de points à changer de classes.

Les principales limites du k -means sont la spécification du nombre de classes, sa sensibilité aux tirages initiaux et aux valeurs extrêmes. L'algorithme n'est pas adapté pour découvrir des clusters qui ne sont pas hyper-elliptiques (ou hyper-sphériques). Par conséquent, il est souvent préférable de transformer l'espace de représentation des données avec des techniques comme l'ACP [122], la classification spectrale [34] ou le scattering operator.

2.4.2.2 Partitionnement par DBSCAN

Density-Based Spatial Clustering of Applications with Noise (DBSCAN) est un algorithme de partitionnement à base de densité proposé par les auteurs de [52] : le but est d'identifier dans l'espace de représentation les zones de forte densité entourées par des zones de faible densité, qui formeront les *clusters*. Son principe est illustré en figure 2.13. Considérant un ensemble de points donnés dans un espace quelconque, les points les plus proches les uns des autres (*i.e.* denses) seront regroupés ensemble.

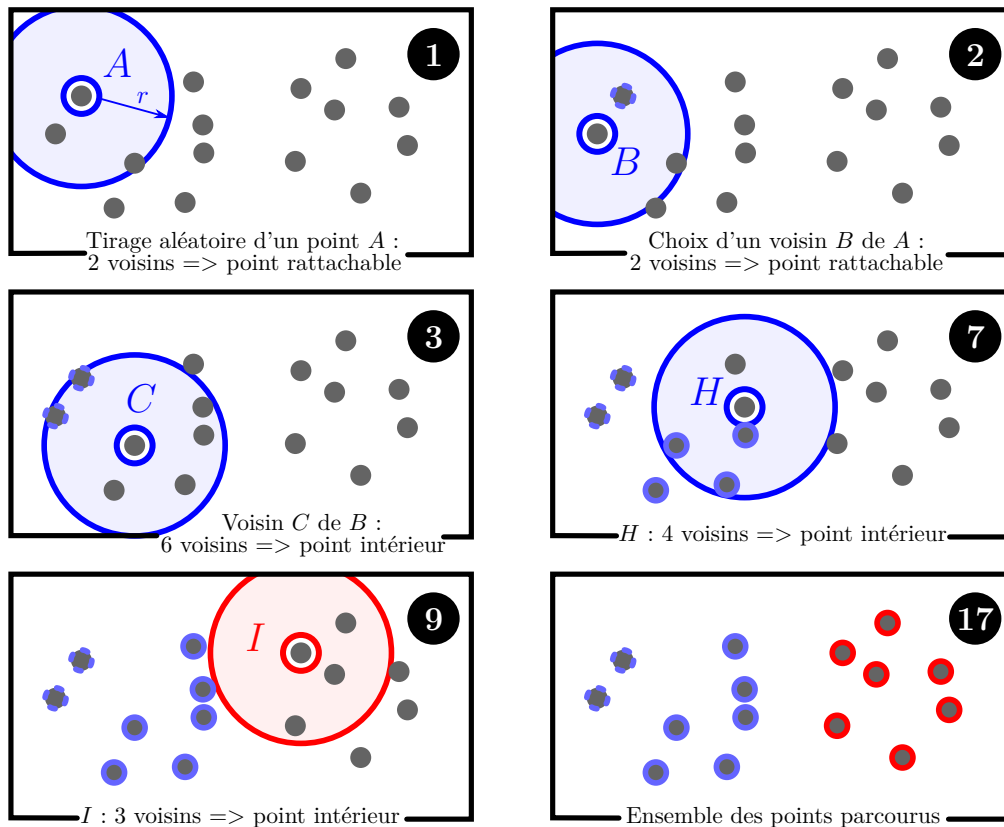


FIGURE 2.13 – Itérations successives de l'algorithme DBSCAN en deux classes (pour un nombre de points minimum pour définir une classe fixé à 3).

En pratique, l'algorithme commence par tirer aléatoirement un point A et recherche tous les points qui se trouvent dans son voisinage suffisamment proche. Le premier paramètre clé à définir est donc la distance radiale r maximale autorisée pour considérer deux points comme voisins. Le second est le nombre de points minimum nécessaire pour créer un *cluster*. En fonction de ce nombre, il y aura trois cas de figure :

1. Le point A n'a aucun voisin et est alors considéré comme **aberrant** ;
2. Le point A a suffisamment de points dans son voisinage, il appartient donc à un *cluster* et est nommé point **intérieur** ;

3. Le point A n'a pas assez de voisins pour être le point intérieur d'un *cluster*, alors il est appelé point **rattachable** et il est attribué au *cluster* de ses voisins (si ses voisins sont des points aberrants, il sera aussi classé comme point aberrant).

Ensuite, si le point A a des voisins, l'algorithme sélectionne l'un d'entre eux et regarde son voisinage. Ainsi, pour chaque point que l'on ajoute au *cluster*, on a une zone de croissance qui va permettre d'étendre le *cluster* de proche en proche. Il convient que plus cette zone (sphère de rayon r) est grande et plus le *cluster* aura de chances de grossir. La notion de voisinage (ou densité) est donc la clé de cette méthode. Contrairement aux k -moyennes, DBSCAN ne nécessite pas de définir le nombre de classes et n'a aucun *a priori* sur la forme de la dispersion des points d'une classes. En revanche, le choix du rayon maximal r impose que les densités des *clusters* aient des échelles similaires. Par exemple, si on a une classe D où les données sont bien denses et une autre E où elles sont bien espacées, le choix d'un r trop faible sera pénalisant pour la classe E (données classées comme aberrantes) alors qu'un rayon r trop grand génèrera probablement qu'un seul *cluster*. Cet inconvénient peut-être pallié en utilisant l'algorithme OPTICS (Ordering points to identify the clustering structure) qui représente des données sous forme de graphe où chaque nœud est rattaché aux nœuds les plus proches de lui.

2.4.2.3 Classification spectrale

Pour pallier aux limites sur la forme des *clusters* des k -moyennes (classes hyperelliptiques) et du DBSCAN (rayon de voisinage constant), il est souvent préférable de transformer l'espace de représentation. Pour regrouper les données les plus similaires entre elles, la classification spectrale (ou *Spectral Clustering*) [71, 163] propose de générer ce nouvel espace à partir du spectre (valeurs propres, vecteurs propres) d'une matrice des similarités des données. Dans la littérature, de nombreuses techniques permettent d'obtenir une mesure de similarité entre deux objets. Le choix de cette mesure dépend essentiellement du domaine d'application et du type de données (variables numériques, binaires, textes). Les auteurs de [34] ont notamment obtenu des résultats prometteurs en appliquant la classification spectrale pour segmenter des IRM de perfusion de prostates. Cette section présente le principe de la classification spectrale ainsi que l'algorithme sélectionné pour réaliser nos simulations en s'appuyant sur les travaux publiés en [71, 163].

a) Principe :

L'idée principale de la Classification Spectrale (CS) est de représenter l'ensemble des données sous la forme d'un graphe pondéré puis de couper ce graphe au niveau des liens les plus faibles. Un graphe est défini par le trio (V, E, W) , où V correspond à ensemble des nœuds du graphe (chaque nœud représente une donnée), E correspond à l'ensemble des couple (v_i, v_j) symbolisant le lien entre deux nœuds et W correspond à la matrice de voisinage. Le coefficient W_{ij} prend comme valeur le poids liant le nœud v_i au nœud v_j . Par définition, la matrice W est symétrique. Un exemple applicatif est illustré en figure 2.14.

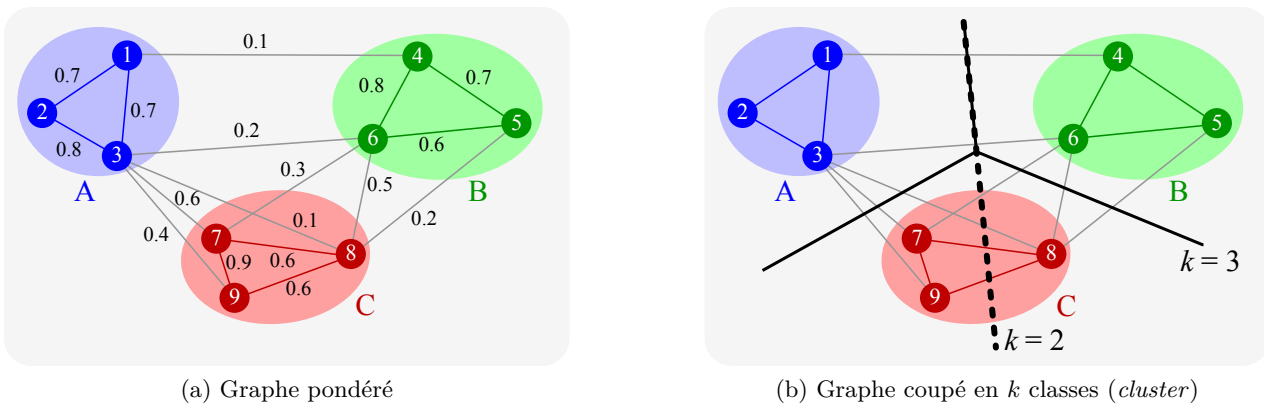


FIGURE 2.14 – Exemple de graphe pondéré puis coupé en deux et trois classes (k).

En pratique, pour un ensemble de nœud donné, il y a plusieurs manières de définir un graphe :

- **Graphe des ε -voisins** : seul les liens entre les nœuds de distance inférieure à ε sont considérés ;
- **Graphe des k -plus proche voisins** : chaque nœud est relié avec les k -plus proches nœuds voisins ;
- **Graphe entièrement connecté** : tous les nœuds sont interconnectés.

Dans nos simulations, nous partons du principe que nous n'avons pas de connaissances *a priori* sur les liens entre les images de thrombi et donc nous utilisons des graphes entièrement connectés. En revanche, le graphe de la figure 2.14.a est un graphe des ε -voisins pour des raisons pratiques ; il est en effet plus simple à représenter qu'un graphe où tous les nœuds sont reliés.

De par la définition du graphe, les notions de similarité et de distance jouent un rôle clé dans les performances de la classification spectrale. Parmi les différentes possibilités qu'offrent la littérature de la classification spectrale, nous avons opté pour la fonction de similitude la plus fréquemment utilisée à savoir la similarité gaussienne W associée à la distance euclidienne d_{euc} :

$$W(I, J) = \exp\left(-\frac{d_{euc}^2(I, J)}{2\sigma^2}\right) \quad (2.12)$$

$$d_{euc}(I, J) = \sqrt{\sum_{i=1}^N (I_i - J_i)^2} \quad (2.13)$$

où I et J sont deux images de la base de données, N le nombre de pixels des images et σ le paramètre qui traduit la dispersion des données. Le choix du paramètre σ est défini à l'annexe E. La matrice de similarité W associe des valeurs comprise entre 0 et 1 à chaque couple de deux nœuds du graphe. Dans l'exemple applicatif, la similarité entre deux nœuds est indiquée par le poids de chaque lien et la matrice de similarité W résultante est affichée au tableau 2.5.c.

L	1	2	3	4	5	6	7	8	9
1	1.5	-0.7	-0.7	-0.1	0	0	0	0	0
2	-0.7	1.5	-0.8	0	0	0	0	0	0
3	-0.7	-0.8	2.5	0	0	-0.2	-0.6	-0.1	-0.4
4	-0.1	0	0	1.6	-0.7	-0.8	0	0	0
5	0	0	0	-0.7	1.5	-0.6	0	-0.2	0
6	0	0	-0.2	-0.8	-0.6	2.4	-0.3	-0.5	0
7	0	0	-0.6	0	0	-0.3	2.4	-0.6	-0.9
8	0	0	-0.1	0	-0.2	-0.5	-0.6	2	-0.6
9	0	0	-0.4	0	0	0	-0.9	-0.6	1.9

(a) Laplacien $L = D - W$

D	1	2	3	4	5	6	7	8	9
1	2.5	0	0	0	0	0	0	0	0
2	0	2.5	0	0	0	0	0	0	0
3	0	0	3.5	0	0	0	0	0	0
4	0	0	0	2.6	0	0	0	0	0
5	0	0	0	0	2.5	0	0	0	0
6	0	0	0	0	0	3.4	0	0	0
7	0	0	0	0	0	0	3.4	0	0
8	0	0	0	0	0	0	0	3	0
9	0	0	0	0	0	0	0	0	2.9

(b) Matrice des degrés D

W	1	2	3	4	5	6	7	8	9
1	1	0.7	0.7	0.1	0	0	0	0	0
2	0.7	1	0.8	0	0	0	0	0	0
3	0.7	0.8	1	0	0	0.2	0.6	0.1	0.4
4	0.1	0	0	1	0.7	0.8	0	0	0
5	0	0	0	0.7	1	0.6	0	0.2	0
6	0	0	0.2	0.8	0.6	1	0.3	0.5	0
7	0	0	0.6	0	0	0.3	1	0.6	0.9
8	0	0	0.1	0	0.2	0.5	0.6	1	0.6
9	0	0	0.4	0	0	0	0.9	0.6	1

(c) Matrice de similarité W

Tableau 2.5 – Calcul du Laplacien non normalisé du graphe de la figure 2.14.a.

Après avoir construit la matrice de similarité qui traduit le poids de la connexion entre chacun des nœuds, on va calculer le degré de connexion de chaque nœud avec l'ensemble des autres nœuds du graphe. Le degré d'un nœud i est défini comme la somme des poids des liens avec les autres nœuds du graphe :

$$D_i = \sum_{j=1}^n W_{ij} \quad (2.14)$$

La matrice D dont la diagonale est constituée des degrés D_i de chaque nœud est appelée matrice des degrés (voir tableau 2.5.b).

Le graphe étant défini, la seconde étape de la Classification Spectrale (CS) consiste à venir couper ce graphe en un certain nombre de groupes k . L'idée intuitive est de le couper au niveau des connexions les plus faibles. Soient X et Y deux ensembles, l'interconnexion Cut (*i.e.* la connexion entre les deux sous-ensembles) est définie comme la somme des poids des liens entre les deux sous-ensembles :

$$Cut(X, Y) = \sum_{i \in X, j \in Y} W_{ij} \quad (2.15)$$

En s'appuyant sur les trois sous-ensembles A , B et C de l'exemple de la figure 2.14, les interconnexions Cut entre ces sous-ensembles valent donc :

$$Cut(A, B) = \sum_{i \in A, j \in B} W_{ij} = 0.3 \quad (2.16)$$

$$Cut(A, C) = \sum_{i \in A, j \in C} W_{ij} = 1.1 \quad (2.17)$$

$$Cut(B, C) = \sum_{i \in A, j \in B} W_{ij} = 1.0 \quad (2.18)$$

$$Cut(A, \bar{A}) = Cut(A, \{B, C\}) = Cut(A, B) + Cut(A, C) = 1.4 \quad (2.19)$$

$$Cut(B, \bar{B}) = Cut(B, \{A, C\}) = Cut(B, A) + Cut(B, C) = 2.1 \quad (2.20)$$

$$Cut(C, \bar{C}) = Cut(C, \{A, B\}) = Cut(C, A) + Cut(C, B) = 1.3 \quad (2.21)$$

$$(2.22)$$

Si l'on considère le sous-ensemble X formé par les points $\{1, 4, 8\}$ et son complémentaire $\bar{X} = \{2, 3, 5, 6, 7, 9\}$, on obtient une interconnexion de 4.9 largement supérieure à celles trouvées précédemment. Pour partitionner le graphe en trois classes, il faut donc trouver les trois sous-ensembles qui minimisent la fonction d'interconnexion (*i.e.* trouver A , B et C dans l'exemple) :

$$J_{kCut}(X_1, \dots, X_k) = \sum_{i=1}^k Cut(X_i, \bar{X}_i) = Cut(A, \bar{A}) + Cut(B, \bar{B}) + Cut(C, \bar{C}) \quad (2.23)$$

En posant $u_i = 1$ si $i \in X_k$ et 0 sinon, cette équation peut également s'écrire sous la forme :

$$J_{kCut}(u_k) = \sum_{i=1}^k u_i^T L u_i \quad (2.24)$$

avec $L = D - W$ et l'exposant T indique la transposée. La matrice L est appelée Laplacien non normalisé de W . Dans notre exemple (voir L dans le tableau 2.5.a), cette fonction d'interconnexion J_{kCut} est bien minimale pour :

$$\begin{aligned} u_A &= [111000000]^T \\ u_B &= [000111000]^T \\ u_C &= [000000111]^T \\ J_{kCut}(A, B, C) &= \sum_{i \in \{A, B, C\}} u_i^T L u_i \end{aligned} \quad (2.25)$$

On peut montrer que la solution pour minimiser la fonction $J_{kCut}(u_k)$ se trouve dans l'espace généré par les vecteurs propres de L associés aux k plus petites valeurs propres de L non nulles. La figure 2.15.b montre la projection des points du graphe de l'exemple dans le plan formé par le deuxième et le troisième vecteur propre (le premier étant associé à 0). Il suffit ensuite d'appliquer un algorithme de classification comme les k -moyennes (décrit en section 2.4.2.1) pour partitionner ces points en k classes dans cet espace simplifié. La figure 2.15.a montre la projection suivant le second vecteur propre dans le cas d'un partitionnement en deux classes. Les résultats du partitionnement sont donnés en figure 2.14b pour 2 et 3 classes. On peut ainsi noter que le choix du nombre de classes est également un paramètre important dans les performances de la classification spectrale.

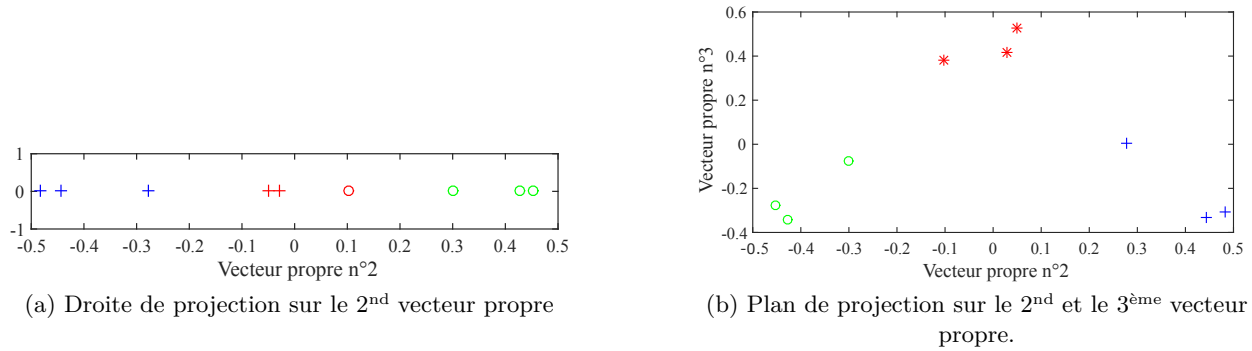


FIGURE 2.15 – Droite et plan de projection des vecteurs propres du Laplacien L du tableau 2.5.a.
 Légende : classe A en bleu, classe B en vert, classe C en rouge, *cluster* 1 avec des ronds, *cluster* 2 avec des croix et *cluster* 3 avec des étoiles.

b) Algorithme utilisé

De nombreux algorithmes existent pour partitionner un graphe en k sous-ensembles. Précédemment, on a vu que l'objectif est de minimiser le poids des connexions entre les sous-ensembles que l'on souhaite dissocier. Une méthode efficace consiste à appliquer l'algorithme de Ng, Jordan et Weiss (NJW) [123] :

1. Construire le graphe et la matrice de similarité W qui va représenter les données, telle qu'elle est définie précédemment dans l'équation 2.12 ;

2. Calculer le Laplacien normalisé L :

$$L = D^{-1/2}(D - W)D^{-1/2} \quad (2.26)$$

où D est la matrice des degrés définie par l'équation 2.14 ;

3. Former la matrice $U = (u_{ij})$ composée des vecteurs propres associés aux k plus petites valeurs propres de L afin d'obtenir une représentation de plus petite dimension ;

4. Normaliser chacune des lignes de U pour former la matrice $T = (t_{ij})$:

$$t_{ij} = \frac{u_{ij}}{\sqrt{\sum_{k=1}^n u_{ik}^2}} ; \quad (2.27)$$

5. Classer chaque ligne de T en k sous-ensemble avec les k -moyennes ou un autre algorithme de classification ;

6. Etiqueter chaque nœud du graphe à partir du résultat de la classification obtenue à l'étape précédente.

2.4.3 Conclusion sur les méthodes de classification

Avec l'explosion quantitative des données numériques (*Big Data*), les techniques d'analyse de données et de classification sont en plein essor. Durant cette section, nous nous sommes focalisés uniquement sur certaines méthodes incontournables en matière de classification que nous avons voulu tester en priorité sur nos données. En supervisé, la classification par SMA est à la fois simple à mettre en place et particulièrement bien adaptée pour classifier les descripteurs du *Scattering Operator* (SO) [22, 161]. Toutefois, pour aller plus loin, il serait intéressant de comparer ses performances à celles des SVM ou du *Deep Learning* (lorsque nous disposerons davantage de données). De plus, nous ne sommes actuellement pas certains que les thrombi, ou leurs images tout du moins, soient classables suivant leurs causes ou la présence d'Embolie Pulmonaire (EP)). C'est pour cette raison que nous proposons également d'utiliser les trois algorithmes non supervisés présentés dans la section 2.4.2. Dans de futurs travaux, nous pourrions aussi envisager d'autres approches de classification comme les arbres de décisions [100, 131], la classification bayésienne [81] ou encore des algorithmes par noyau (SVM [93], ACP [87], k -moyennes [72]).

2.5 Conclusion sur l'état de l'art et le choix des approches développées

Dans ce chapitre, nous avons pu constater que la littérature sur la caractérisation de tissus humains (mammaires, hépatiques, prostatiques, thyroïdiens et salivaires) par imagerie acoustique (échographie ou élastographie) est beaucoup plus conséquente que celle sur les TVP. Parmi les quatre types de prétraitement énoncés pour améliorer la qualité de l'image, nous considérerons par la suite uniquement l'EH des échographies. L'objectif de cette EH est de normaliser les échographies par rapport aux changements de paramètres d'acquisition inter-patients (comme le gain). La réduction du bruit étant déjà pris en compte par le système Toshiba Aplio 500, nous avons préféré ne pas ajouter de prétraitements supplémentaires (qui pourraient entraîner des effets contraires à ceux souhaités).

En ce qui concerne la segmentation, celle-ci est réalisée soit manuellement avec validation par un expert médical, soit à partir des ellipses entourant les thrombi et dessinés par les experts médicaux. La différence de contraste entre le thrombus et son voisinage (parois, artères, muscles) compliquerait l'automatisation de cette tâche et ne présente pas de réel intérêt pour la caractérisation du thrombus (la forme du thrombus est souvent elliptique en transversale, cylindrique en longitudinale, l'ellipse est systématiquement dessinée pour l'élastographie).

Parmi les différentes familles de descripteurs, nous avons envisagé celles des ondelettes et celles des statistiques qui s'intéressent respectivement à la texture de l'image et à sa distribution. Les descripteurs morphologiques ne paraissent pas forcément adaptés à la problématique et il n'existe pas, à ce jour, de modèle mathématique capable de traduire la structure et/ou l'évolution d'un thrombus. Ensuite, pour observer si nos données peuvent être partitionnées et si les *clusters* obtenus peuvent être reliés aux informations relatives à la maladie, nous envisageons d'abord une méthode de visualisation de la surface des données (t-SNE), une méthode supervisée (l'SMA) et trois algorithmes non supervisés (les k -moyennes, DBSCAN et la classification spectrale). Les trois prochains chapitres sont respectivement dédiés à la préparation des données, à l'extraction des descripteurs et aux résultats de classification.

Chapitre 3

Base de données et prétraitements

Glossaire du chapitre 3

CLAHE Egalisation d'histogramme locale : *Contrast Limited Adaptative Histogram Equalization*

DICOM Format standard pour la gestion des images biomédicales (*Digital Imaging and Communications in Medicine*)

DTHI Differential Tissue Harmonic Imaging : échographie des tissus humains basée sur l'analyse de la différence des harmoniques de deux fréquences

EP Embolie Pulmonaire

ID Identifiant

IHM Interface Homme-Machine

RGB Codage des couleurs en rouge, vert et bleu (*Red, Green, Blue* en anglais)

ROI Région d'intérêt (*Region Of Interest*)

THI Tissue Harmonic Imaging : échographie des tissus humains formée à partir des harmoniques des ultrasons réfléchis

Introduction

Les deux chapitres précédents ont présenté la problématique du projet et fait le point sur les travaux existants dans le contexte de la thèse. L'objectif majeur de la thèse est de caractériser la structure du thrombus veineux par échographie et élastographie afin de la corréliser avec l'épidémiologie du patient (facteurs de risque, embolie pulmonaire). Dans la littérature, nous avons pu découvrir un certain nombre de méthodes qui pourraient nous permettre d'approcher nos objectifs et notamment pour calculer des descripteurs de nos images (ondelettes, statistiques) puis les classer (classifieur supervisé basé sur l'Analyse en Composante Principale, k -moyennes, classification spectrale). Néanmoins, ces méthodes ne peuvent pas être appliquées directement sur les images de nos bases de données (thrombi et glandes) construites avec le système Toshiba Aplio 500. Ce chapitre dresse l'ensemble des étapes développées pour préparer nos données. La section 3.1 décrit les bases en question et les besoins en matière de préparation. La section 3.2 est consacrée au premier bloc de l'analyse de l'information (voir la chaîne de traitement de la figure 2.2), nommé « prétraitements ». Cette section montre notamment comment les informations (paramètres d'acquisition, épidémiologie du patient) et les différentes images (échographie, élastographie et carte de propagation) sont récupérées des fichiers des bases de données. La section 3.3 correspond au deuxième bloc de l'analyse de l'information qui regroupe la segmentation des images et l'extraction des images utilisées pour le calcul des descripteurs.

3.1 Description de la base de données

Cette première section présente les trois bases utilisées par la suite lors des simulations : thrombi, glandes et textures. La base de thrombi veineux (*cf.* sous-section 3.1.1) est constituée des données obtenues avec le protocole 3 (*cf.* sous-section 1.3.2.3) qui correspond au protocole le plus récent, avec le plus grand nombre d’images et avec la meilleure qualité d’élastographie. Comme évoqué au chapitre 1, le service de rhumatologie du CHRU de Brest a collecté des acquisitions de glandes salivaires avec le même système que celles de thrombi (Toshiba Aplio 500). A l’aide de cette base résumée en sous-section 3.1.2, nous allons évaluer les performances des approches développées sur le même type de données que celles de thrombi. La troisième base est une base d’images synthétiques contenant dix types de textures et est utilisée pour évaluer et paramétrer les différentes approches d’extraction de descripteurs et de classification. Elle est présentée en section 5.1.

3.1.1 Base de thrombus

La base de thrombi veineux collecté avec le protocole 3 (*cf.* sous-section 1.3.2.3) comprend trois types d’images : échographies, élastographies et cartes de propagation des ondes de cisaillement (voir sous-section 1.2.3.2). Les thrombi sont également observés selon deux plans de coupe : transversal et longitudinal. Les images sont des captures d’écran du système Toshiba Aplio 500 et sont scindées en deux parties distinctes : l’élastographie à gauche et l’échographie ou la carte de propagation à droite. Lors des acquisitions, le praticien dessine une ellipse (ou plusieurs en longitudinal) au niveau du thrombus pour indiquer sa présence et estimer la vitesse moyenne des ondes de cisaillement à l’intérieur de l’ellipse (mesure appelée élastométrie). Pour un patient donné, il y a donc aux moins deux acquisitions (transversale et longitudinale) comprenant chacune une ou plusieurs captures (exemple à la figure 3.1). Lors de la préparation des données, il faudra donc rassembler les images d’une même acquisition et séparer les parties gauche et droite de chaque image. De plus, en extrayant la position de l’ellipse, il sera donc possible d’extraire des imagerie (*i.e.* une petite image) comportant uniquement du thrombus sans réaliser de segmentation proprement-dite.

Dans le tableau 1.5 du chapitre 1, nous avons pu constater que les acquisitions n’ont pas toutes été obtenues avec les mêmes paramètres : les modes fondamental et Tissue Harmonic Imaging (THI) et un gain variant entre 67 et 75 *dB* (voir la section 1.2.2 pour plus de détails). Afin d’observer l’impact de ces modifications de paramètres lors des simulations, ces paramètres devront donc être récupérés lors de la phase de préparation des données. La sous-section 3.2.3 propose de réaliser une égalisation d’histogramme pour limiter l’influence des variations du gain (qui permet d’augmenter l’intensité des échos). Enfin, comme nous pouvons l’observer sur la figure 3.1, l’élastographie, superposée à l’échographie (parfois avec transparence comme sur cet exemple), est une carte couleur codée en RGB (*Red, Green, Blue*). Pour l’exploiter, une étape de conversion en carte de vitesse à partir de l’échelle des couleurs est donc nécessaire. Les prochaines sections de ce chapitre illustrent les différentes étapes de la préparation des données de thrombi (environ 250 acquisitions répartis sur 36 patients), mais aussi de glandes salivaires.

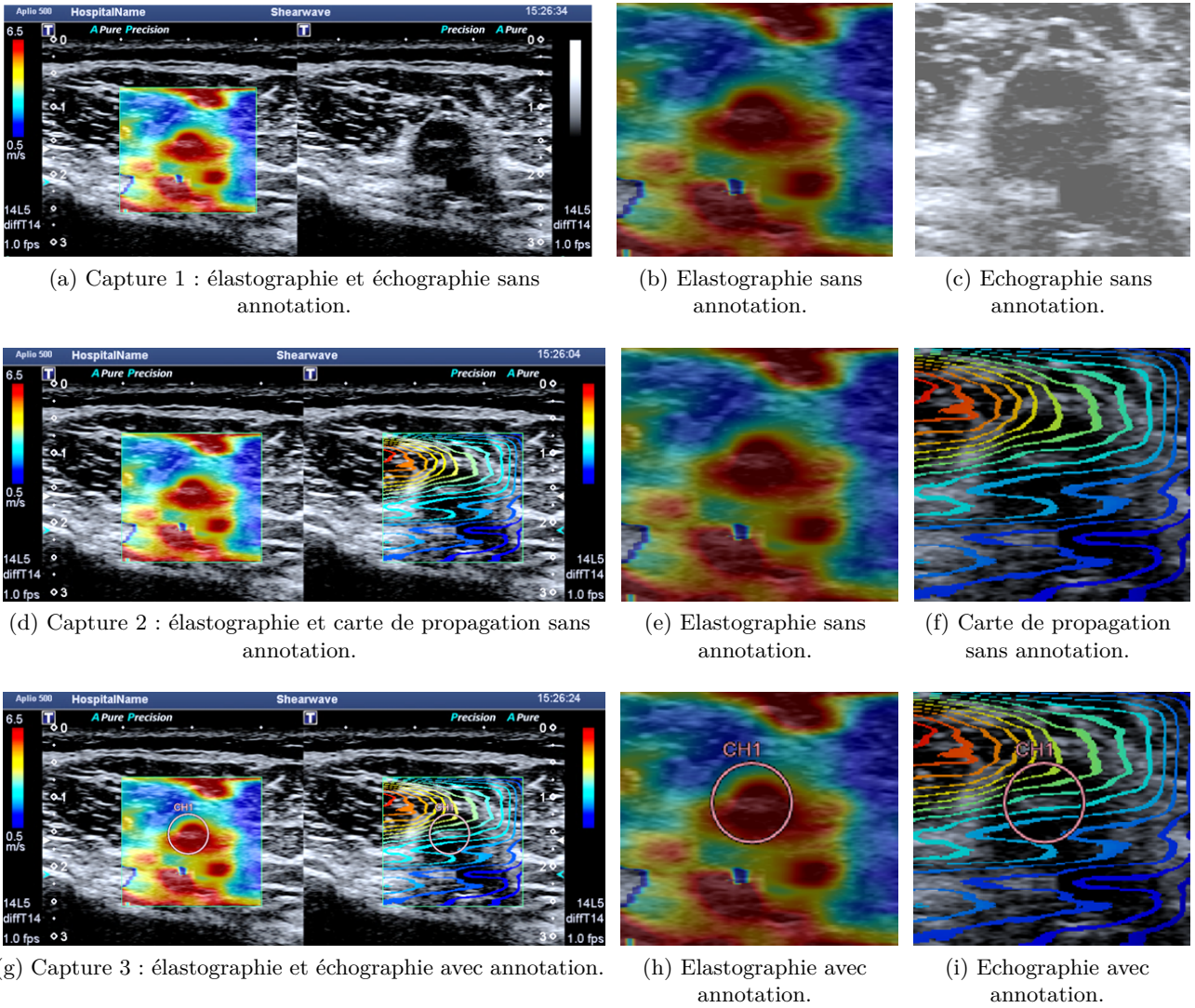
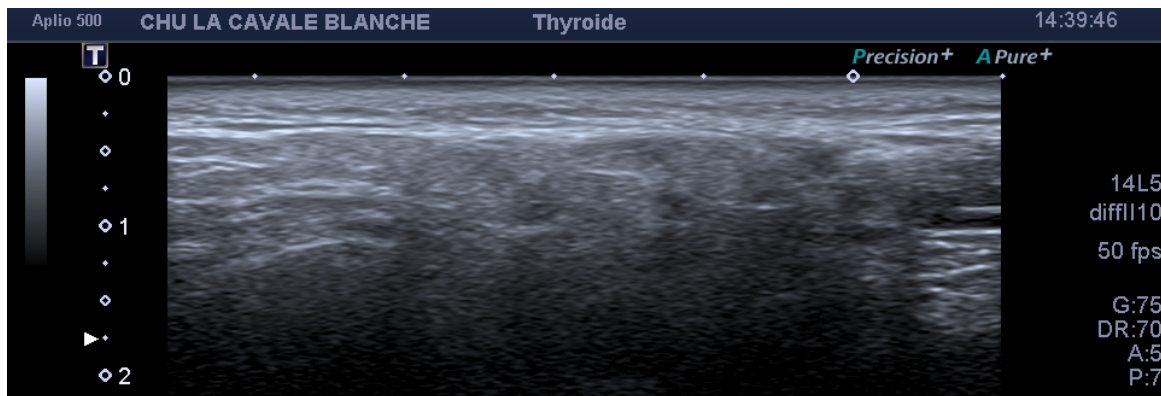


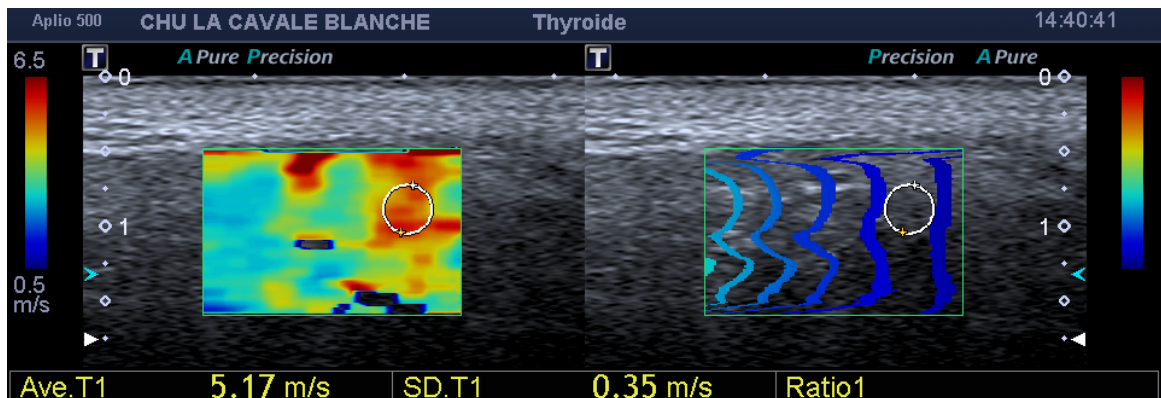
FIGURE 3.1 – Extraction des différentes images pour une acquisition (plusieurs captures d'écran) : échographie, élastographie et carte de propagation avec ou sans annotation.

3.1.2 Base de glandes salivaires

La base de glandes salivaires, également collectée avec le système Toshiba Aplio 500, contient les trois mêmes types d'images, illustrés en figure 3.2, que pour les thrombi (échographies, élastographies, cartes de propagation). Pour chaque patient, le praticien examine chacune des six glandes principales : parotides, submandibulaires et sublinguales gauche et droite (*cf.* section 1.3.3). Toutefois, les glandes sublinguales ne seront pas considérées dans cette étude car la très grande majorité d'entre elles sont saines et les élastographies sont souvent de mauvaises qualités (nombreuses zones sans valeurs). Pour les quatre glandes étudiées, il y a au moins une capture de l'échogénéicité de la glande (figure 3.2.a) et une autre de la dureté (figure 3.2.b). Comme pour les thrombi, une mesure d'élastométrie (vitesse moyenne dans une région d'intérêt) est effectuée sur une portion de la glande ; il peut donc y avoir deux captures pour une acquisition d'élastographie (avec et sans l'ellipse). L'ellipse étant petite par rapport à la taille des glandes, les images de glandes salivaires, et notamment les échographies, doivent donc être segmentées pour en extraire des imageriettes (*cf.* section 3.3.2).



(a) Capture 1 : échographie d'une glande parotide.



(b) Elastographie sans annotation.

FIGURE 3.2 – Extraction des différentes images pour une acquisition (plusieurs captures d'écran) : échographie, élastographie et carte de propagation avec ou sans annotation.

Au final, sur les 37 patients de l'étude, soixante-dix acquisitions ont été collectées pour chaque mode d'imagerie (échographie et élastographie) et chaque type de glandes (parotides et submandibulaires). Après consultation avec les experts médicaux, environ soixante pourcent des glandes sont considérées comme normales et donc quarante pourcent comme atteintes du syndrome de Gougerot-Sjörden (section 1.3.3). Parmi les glandes malades, il est à noter que certaines n'ont pas le même grade d'avancement (grades allant de 1 à 3 en fonction de l'avancée de la maladie). Il est donc possible que les glandes atteintes du syndrome de grade 1 (grade le moins avancé) se confondent plus facilement avec les glandes normales. Enfin, de manière similaire aux examens des veines, le praticien peut ajuster certains paramètres d'acquisition comme le gain ; les échographies seront également soumises à une égalisation d'histogramme comme nous le verrons en section 3.2.3.

Conclusion sur les bases de données

Les acquisitions de thrombi et de glandes salivaires ont été recueillies avec le système Toshiba Aplio 500 qui ne permet pas d'accéder aux données brutes mais à des captures d'écran du système. Une capture peut représenter seulement l'échographie (glandes salivaires) ou alors être séparée en deux parties pour afficher deux modes d'imagerie parmi : l'échographie, l'élastographie et la carte de propagation. Une même acquisition peut être composée de plusieurs captures (avec ou sans annotations). Avant de pouvoir appliquer les différentes approches d'extraction de paramètres aux images de nos bases, un certain nombre de prétraitements sont à effectuer en amont : lecture et extraction des images, identification des paramètres d'acquisition, conversion des élastographies, localisation des régions d'intérêt et extraction d'images de ces régions.

3.2 Préparation des données

Cette section présente l'ensemble des étapes de prétraitement effectué avant la localisation des régions d'intérêt et l'extraction des images. La première étape de la préparation des données consiste à lire les images, à identifier leurs types présents dans chacune des captures (échographie, élastographie et carte de propagation) et à séparer les deux parties de chaque capture. Il faut également récupérer les paramètres d'acquisition pour pouvoir estimer par la suite leur influence sur les résultats de classification. Ensuite, les images doivent être également regroupées par acquisition. L'ensemble de cette étape fait l'objet des sections 3.2.1 et 3.2.2 pour, respectivement, les images de thrombi et celles de glandes. Les deux étapes suivantes, décrites en section 3.2.3 et 3.2.4, consistent à normaliser les échographies (égalisation d'histogramme) et convertir les élastographies en carte de vitesse.

3.2.1 Cas des thrombi

En ce qui concerne la base de thrombus, deux scripts principaux ont été créés pour réaliser cette première étape de préparation des données. Le premier script génère l'Interface Homme-Machine (IHM) *Lectio*, présentée en sous-section 3.2.1.1, qui s'occupe de lire les données, de les identifier et de récolter les informations importantes sur l'acquisition.

3.2.1.1 Lecture et conversion

Les captures d'écran, collectées lors des examens médicaux, sont exportées au format DICOM (*Digital Imaging and Communications in Medicine*) : elles contiennent l'image et des métadonnées, c'est-à-dire des données sur l'acquisition (système, sonde, hôpital, patient, date de l'examen, *etc.*). Toutefois, certaines informations ne sont donc pas disponibles dans les métadonnées mais sont accessibles directement sur l'image elle-même :

- L'élastométrie : moyenne et écart-type des vitesses des ondes de cisaillement dans la zone d'intérêt) ;
- Les initiales de l'expert médical et l'identifiant de l'acquisition (« Measure ID ») ;
- La présence d'un thrombus : l'expert médical précise si la veine est libre en annotant l'image ;
- La présence de l'ellipse autour du thrombus ;
- Les réglages utilisés pour obtenir l'image échographique (gain, dynamique, fréquence, *etc.*) ;
- Le type des deux parties présentes à l'écran (échographie, élastographie, propagation des ondes de cisaillement).

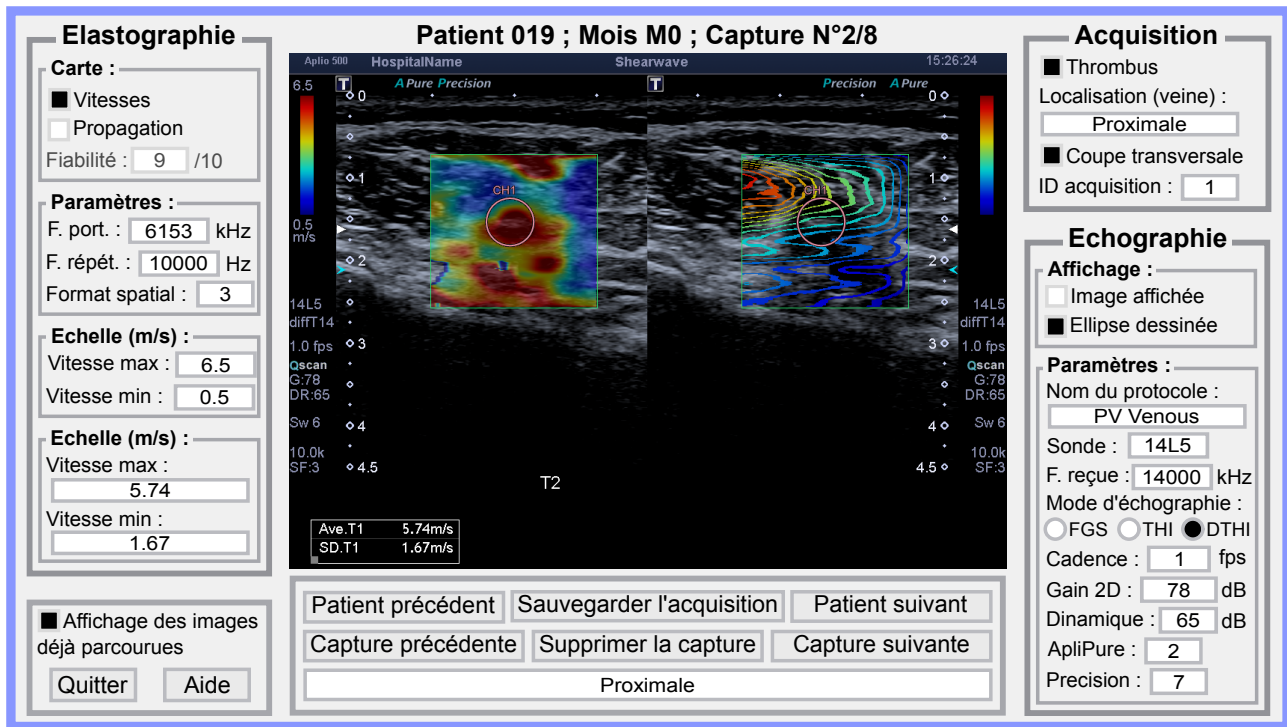


FIGURE 3.3 – IHM *Lectio* pour lire les fichiers DICOM et récupérer les images, les paramètres d'acquisition et les informations relatives aux patients.

Pour récupérer les informations disponibles dans les métadonnées et celles affichées directement sur l'image, une IHM nommée *Lectio* a été créée ; elle est visible en figure 3.3. Lors de l'utilisation de l'IHM *Lectio*, une fois les différents champs remplis, l'utilisateur sauvegarde les changements et passe au prochain fichier. L'IHM se charge également d'aller récupérer les données médicales sur le patient qui se trouvent dans un tableur : classe (cf. tableau 1.5), identifiant du patient, âge, sexe, antécédents, localisation de la tête du thrombus, cancer, facteur génétique, tabac et Embolie Pulmonaire (EP). Lors de la sauvegarde, l'IHM enregistre, dans un dossier que nous allons appeler *Dir 1*, l'image et l'ensemble des informations extraites dans une même structure. L'utilisateur peut alors passer au fichier suivant et répéter l'opération jusqu'à avoir parcouru la totalité des captures d'écran de la base.

3.2.1.2 Découpage, résolution et type d'images

Le deuxième script principal a pour objectif d'extraire les images présentes dans les captures (comme sur la figure 3.1) et rassembler les données d'une même acquisition. Pour ce faire, ce script parcourt l'ensemble des structures du dossier *Dir 1* issues de l'IHM *Lectio*. Pour une structure donnée, une fonction permet d'extraire les deux parties de la capture d'écran en s'appuyant sur les métadonnées : elles donnent le coin en haut à gauche, le coin en bas à droite et le centre de profondeur zéro des images de gauche et de droite (ces points clés sont indiqués par des étoiles sur la figure 3.4). Le rognage est effectué à partir de ces coordonnées et de telle manière que les graduations soient exclues des images extraites (cf. cadre orange figure 3.4). Toutefois, juste avant le rognage, un seuillage sur les couleurs des graduations est effectué pour obtenir leurs coordonnées et ainsi la résolution en profondeur de l'échographie. Ensuite, les deux images extraites sont identifiées (type d'images, avec ou sans ellipse) à partir des informations recueillies par l'IHM *Lectio* et stockées dans une nouvelle structure dans un nouveau dossier *Dir 2*. Dans cette même structure sont également ajoutées toutes les informations relatives à l'acquisition et au patient.

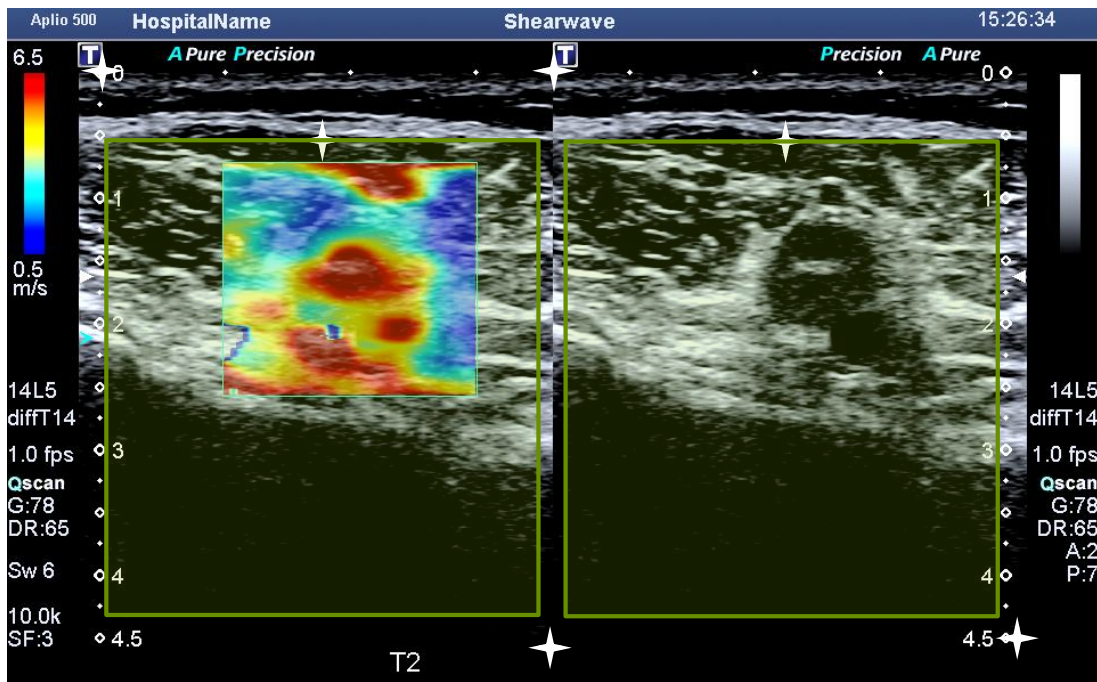


FIGURE 3.4 – Rognage des images à partir des métadonnées (image acquise avec le protocole 1) : les métadonnées donnent les coordonnées des points représentées par des étoiles en blanc. Les deux parties de l'image (élastographie et échographie) sont rognées de telle sorte qu'elles aient la même taille et que les graduations ne soient plus visibles (*cf.* images encadrées).

Ensuite, le script passe à la capture suivante (stockée dans une structure du dossier *Dir 1*). Si elle correspond à la même acquisition (même *Measure ID*), les deux images (gauche et droite) sont extraites et ajoutées à la structure de l'acquisition. Dans le cas contraire, elles seront enregistrées plus tard dans une nouvelle structure. Une fois l'ensemble des images d'une acquisition extraites, un second rognage à la taille de l'élastographie est effectué sur chacune des images : en soustrayant l'échographie par l'élastographie, on obtient un masque de l'élastographie et donc sa position. Le résultat de ce second rognage est illustré sur l'exemple de la figure 3.1. La section 3.2.4 montre comment l'élastographie est alors convertie en carte de vitesse et la section 3.2.3 comment est extraite la position de l'ellipse dans ses images. Enfin, l'ensemble des images extraites, la carte de vitesses, les coordonnées de l'ellipse et l'ensemble des informations relatives à l'acquisition sont enregistrés dans un nouveau dossier *Dir 2* sous un identifiant unique : Px_My_Dz où x représente l'identifiant (ID) patient, y l'âge du thrombus en mois (0, 1 ou 3) et l'ID de l'acquisition. Le script passe ensuite à la prochaine acquisition et ainsi de suite jusqu'à avoir parcouru l'ensemble des fichiers du dossier *Dir 1*.

3.2.2 Cas des glandes salivaires

En ce qui concerne les glandes salivaires, les deux étapes de préparation décrites dans la section précédente ont été simplifiées. Chaque capture a été renommée de manière unique en fonction de :

- ID patient (*e.g.* $P01$) ;
- ID de l'acquisition (*e.g.* $D001$) ;
- Glande observée : parotide (PA), submandibulaire (SM), sublinguale (SL) ;
- Côté observé : gauche (G) ou droite (D) ce qui donne PAG pour une parotide gauche ;
- Etat de la glande : normale (N) ou Gougerot (G) ;
- Type d'image : échographie ($Echo$) ou élastographie ($Elas$) ;
- Présence de l'ellipse sur l'élastographie (*val* avec, *map* sans).

Par exemple, l'échographie d'une parotide normale droite du patient 11 est identifiée par *P11_D003_PAD_N_Echo* alors que l'élastographie de cette même glande, avec l'ellipse indiquée, est renommée par *P11_D003_PAD_N_Elas_val*. Ensuite comme pour les images de thrombus, ces images sont lues, rognées à partir des métadonnées et stockées dans une structure propre à chaque acquisition. Les élastographies avec et sans ellipse sont stockées dans la même structure. Ensuite, l'élastographie (sans ellipse si possible) est convertie en carte de vitesse et ajoutée à la structure. En section 3.3.2 et 3.3.3, nous nous intéresserons à la segmentation des glandes en échographie et à l'extraction des imagettes.

3.2.3 Egalisation d'histogramme des échographies

Pour un mode d'échographie donné (fondamental, THI ou DTHI), le gain (*cf.* section 1.3.1.3) est le principal paramètre variable entre les acquisitions des différents patients. L'augmentation du gain permet d'amplifier le signal reçu et donc de mieux visualiser les structures peu échogènes. En contrepartie, certaines structures très échogènes peuvent être saturées. Pour uniformiser au mieux les échographies obtenues avec des gains différents, nous proposons d'appliquer une égalisation d'histogramme sur chaque échographie, thrombus ou glande. Cette méthode consiste à distribuer de façon uniforme les niveaux de gris de l'image sur toute la plage possible (de 0 à 255), comme l'illustre la figure 3.5.b. On peut voir que l'égalisation d'histogramme a amélioré le contraste général en éclaircissant l'image mais que les zones les plus échogènes ont été saturées (parois de la veine).

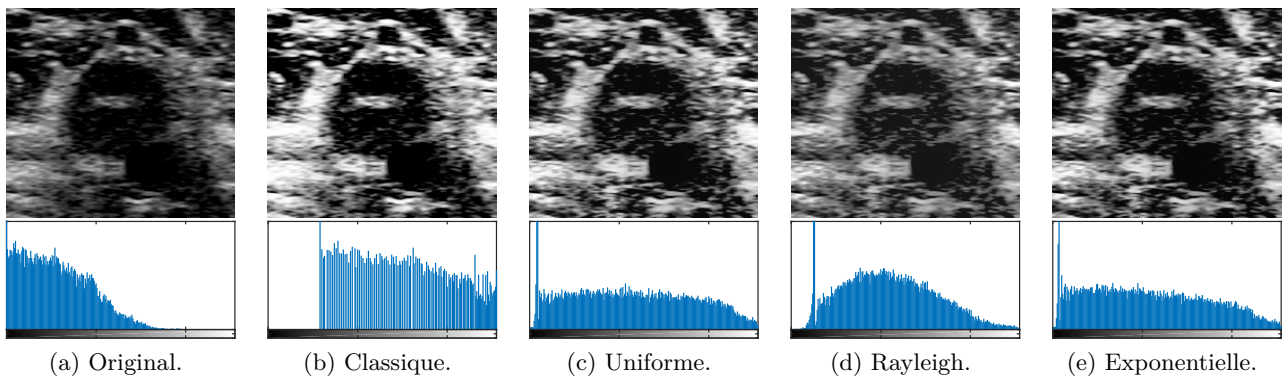


FIGURE 3.5 – Echographies et histogrammes avant et après égalisation d'histogramme classique ou Contrast Limited Adaptive Histogram Equalization (CLAHE) [166] avec trois types de distribution : uniforme, Rayleigh ou exponentielle.

Pour améliorer le contraste local et éviter ces phénomènes de saturation, les auteurs de [166] proposent d'effectuer une égalisation d'histogramme locale. Au lieu de transformer l'histogramme sur l'ensemble de l'image, la méthode consiste à égaliser l'histogramme sur une fenêtre glissante qui va parcourir l'ensemble de l'image. Un seuil est également fixé pour éviter d'amplifier fortement des pixels de bruits (notamment dans les régions homogènes) et donc de dénaturer l'image. Cette méthode nommée *Contrast Limited Adaptive Histogram Equalization* (CLAHE) permet donc d'améliorer le contraste localement tout en évitant des phénomènes de saturation. La loi de distribution peut être uniforme (comme pour l'égalisation d'histogramme classique) mais peut aussi suivre la loi de Rayleigh ou la loi exponentielle. On peut observer en figure 3.5 que les méthodes locales (CLAHE) donnent de meilleurs résultats que la méthode classique : les parois ne sont plus saturées et le contraste au niveau du thrombus est amélioré. Par contre, il est plus compliqué de déterminer quelle est la meilleure loi de distribution dans ce cas particulier.

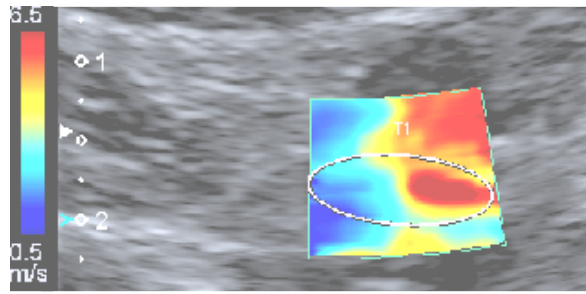
3.2.4 Conversion des élastographies en cartes de vitesses

Comme il est possible de le voir sur les figures 3.6.a et 3.6.c, l'élastographie est affichée, soit directement (figure 3.6.a), soit en transparence (figure 3.6.c) sur l'échographie. Dans le premier cas, le rognage décrit à la section 3.2.1.2 suffit pour n'obtenir que l'élastographie. Dans le second cas, une

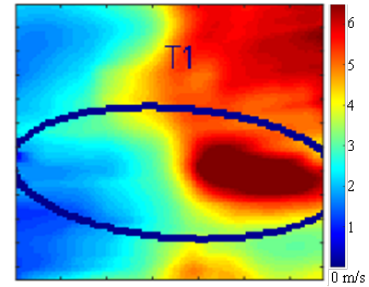
étape préliminaire est nécessaire afin de supprimer l'échographie. Pour ce faire, il est donc nécessaire de disposer de l'échographie I_{echo} (illustré en figure 3.6.d) et du coefficient de transparence γ (qui est fixé à 0 ou 0.5 par le système Toshiba Aplio 500). Ensuite, l'élastographie seule I_{elasto} à partir de l'image de départ $I_{echo+elasto}$ s'obtient en appliquant la formule suivante :

$$I_{elasto} = (1 - \gamma)I_{echo+elasto} + \gamma I_{echo} \quad (3.1)$$

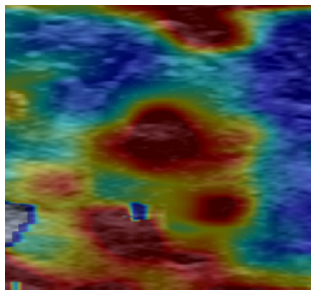
Dans les deux cas, l'élastographie est une image couleur en RGB qui, pour être utilisée, est convertie en carte de vitesse. L'échelle des vitesses affichées sur la bordure gauche de la capture d'écran (voir figures 3.1 ou 3.6.a) permet d'estimer les vitesses correspondant à chaque trio de valeurs RGB. Deux exemples de conversion en carte de vitesses sont visibles dans les figures 3.6.b et 3.6.e.



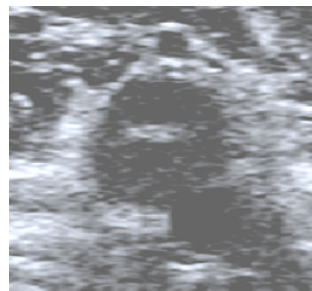
(a) Elastographie couleur sans transparence.



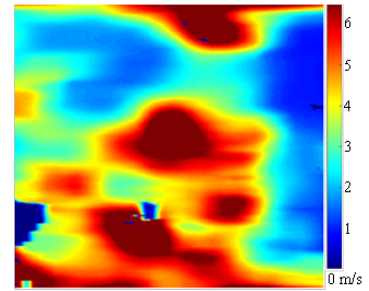
(b) Elastographie (a) convertie en carte de vitesses.



(c) Elastographie couleur en transparence sur l'échographie.



(d) Echographie.



(e) Elastographie (c) convertie en carte de vitesse.

FIGURE 3.6 – Conversion de l'élastographie couleur (chaque pixel correspond à un niveau de couleur rouge, vert et bleu) en carte de vitesses (chaque pixel correspond à une vitesse en m/s) : dans le cas où l'élastographie est en transparence sur l'échographie, une étape supplémentaire est nécessaire afin de supprimer l'échographie en transparence.

3.2.5 Conclusion sur la préparation des données

Au départ, nous disposons de captures d'écran du système Toshiba Aplio 500 enregistré au format DICOM. Ce format permet de contenir, en plus de l'image, certaines informations (appelées métadonnées) sur les conditions d'acquisition et le patient. Néanmoins, certains paramètres d'acquisition, l'élastométrie et l'épidémiologie des patients ne sont pas présents dans les métadonnées. J'ai donc développé l'IHM *Lectio* afin de récupérer ces données manquantes et les stocker, avec l'image, dans une nouvelle structure. Dans cette section, nous avons également constaté qu'une capture est composée de deux types d'images (sauf pour les échographies des glandes salivaires) : élastographie et échographie, ou élastographie et carte de propagation. De plus, lors de l'examen, l'expert médical place une ellipse sur l'image pour indiquer la position du thrombus et calculer l'élastométrie. A l'issue des étapes présentées dans cette section, nous disposons d'une structure par acquisition qui contient les différents types d'image de l'acquisition, l'échographie après égalisation d'histogramme, l'élastographie convertie en carte de vitesse et la position de l'ellipse. Les images sont donc prêtes (ou presque, pour les échographies) pour passer à l'étape d'extraction d'images.

3.3 Segmentation et extraction d'images de la région d'intérêt

Dans cette section, nous abordons l'étape d'extraction des imagerie. L'objectif est d'extraire une portion d'image la plus significative possible de la région d'intérêt (le thrombus), de taille adaptée pour limiter et simplifier les calculs lors de l'extraction des descripteurs (*cf.* chapitre 4). Cette étape a évolué au cours de la thèse pour s'adapter à l'évolution du protocole de collecte des données, comme le montre l'annexe B. La veine observée (poplitée, fémorale, *etc.*) et l'anatomie du patient différent (taille, profondeur, position). Par conséquent, dans un premier temps, seule une imagerie était extraite à l'intérieur de l'ellipse, puis elle était redimensionnée pour que toutes les imagerie extraites aient la même taille. Ces procédures sont expliquées en annexe B (sections B.3 et B.4). Après réflexion, nous avons remarqué que les parois pouvaient également être source d'information (parois marquées en cas d'inflammation sur l'élastographie) et qu'il serait peut-être judicieux d'extraire une imagerie de référence (tissu alentour) pour pouvoir prendre en compte la notion de contraste entre le thrombus et son voisinage. La section 3.3.1 présente une application pour segmenter manuellement et rapidement les échographies de thrombus puis pour extraire automatiquement les imagerie. La section 3.3.2 aborde ces deux mêmes problématiques pour les échographies des glandes salivaires. Enfin, pour traiter les élastographies, une autre application, détaillée en section 3.3.3, a été développée pour déterminer manuellement la position des imagerie car les trous (régions sans valeurs) des élastographies compliquent leur extraction automatique.

3.3.1 Echographies des thrombi

L'extraction des imagerie sur les échographies se déroule en trois étapes : extraction d'une image centrée autour de l'ellipse (section 3.3.1.1), segmentation de cette image et extraction automatique d'imagerie à partir de cette segmentation.

3.3.1.1 Détection et position de la région d'intérêt

Sur la capture d'écran contenant l'échographie et l'élastographie, l'ellipse dessinée par les praticiens sur le thrombus constitue notre région d'intérêt région d'intérêt ROI (*Region of Interest*). Elle possède des couleurs spécifiques qui vont permettre de l'identifier sur l'image couleur. En fixant à 1 tous les pixels de l'image qui ont ces couleurs et à 0 les autres pixels, on obtient une image binarisée qui représente le contour de l'ellipse (et quelques pixels de l'image ou des autres annotations). Ensuite, en appliquant un filtre de dilatation des pixels (pour fermer le contour s'il ne l'est pas), un autre filtre de remplissage et un d'érosion, nous obtenons, en règle générale, un parfait masque de l'ellipse. L'annexe B.3.1 revient sur les cas particuliers qui peuvent poser problème pour segmenter l'ellipse (ellipse tronquée, plusieurs ellipses sur l'image) et les solutions apportées. Une fois le masque obtenu, il reste à estimer les caractéristiques (centre, petit axe, grand axe et orientation) et la position de l'ellipse (*i.e.* du masque), ce qui peut se faire facilement à partir de fonctions existantes (comme *regionprops*).

3.3.1.2 Segmentation de la région d'intérêt

L'objectif ici est de créer un outil simple et rapide qui permet de diviser l'échographie de thrombus en trois régions distinctes : thrombus, voisinage (artère, muscle) et mixte (thrombus, parois et voisinage). Pour ce faire, l'idée est d'extraire une image carrée de taille fixe (même taille pour chaque acquisition) et partitionnée en un nombre fixe de cases interactives. Ces cases changent de teintes en fonction du nombre de clics :

- Aucun clic : la portion d'image de la case n'est pas conservée ;
- Un clic : cette portion contient uniquement du thrombus ;
- Deux clics : cette case couvre une grande partie du thrombus mais également du tissu extérieur ;
- Trois clics : très peu ou aucun morceau de thrombus n'est présent dans le masque.

Ensuite, une fois que l'utilisateur a segmenté l'image en interagissant avec les différentes cases, un masque est créé et les imagerie sont extraites et étiquetées en fonction des nombres de clics. La figure 3.7 illustre la réalisation de cet outil de segmentation.

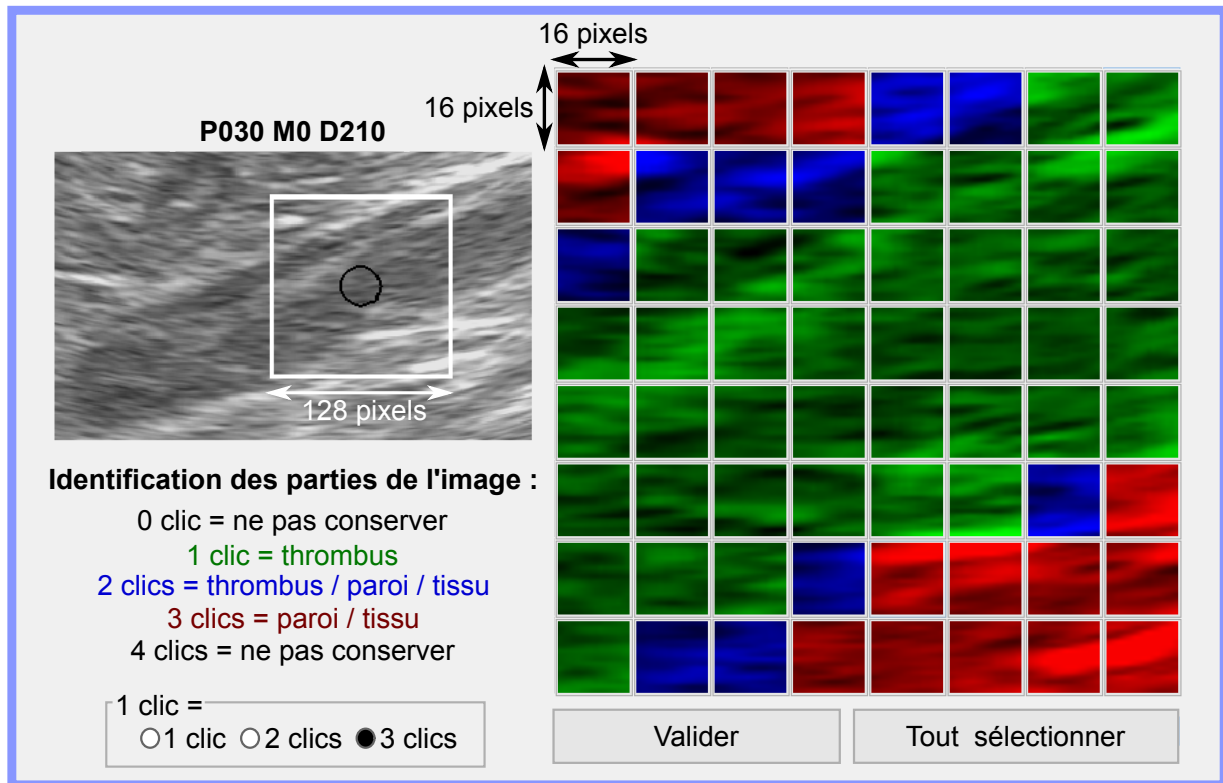


FIGURE 3.7 – IHM *Mosaic* pour segmenter les échographies de thrombus. Chaque case est teintée en fonction du nombre de clics de l'utilisateur sur celle-ci. Chaque teinte correspond à une étiquette parmi « Thrombus », « Mixte » ou « Tissu ». Il est possible de cliquer sur toutes les cases à la fois en appuyant sur le bouton « Select All ».

En pratique, pour réaliser l'IHM *Mosaic* de la figure 3.7, il y a trois paramètres à définir : la taille de la fenêtre de segmentation, sa position et la taille des cases de la fenêtre. Au vu de la taille des images échographiques (entre 150×150 et 300×300 pixels), nous avons opté pour une fenêtre de taille 128×128 pixels afin d'être sûr qu'une telle image puisse être extraite. De plus, le but étant de caractériser le thrombus, cette fenêtre a donc été centrée sur l'ellipse (qui indique la position du thrombus). Ensuite, pour la taille des cases interactives, il fallait trouver le bon compromis entre le nombre de cases à interagir (qui ne doit pas être trop important pour des questions de rapidité) et la taille des imagerie que nous souhaitons extraire. Or, par rapport à la taille des thrombi et aux contraintes du scattering operator (voir section 4.1) avec lequel les calculs sont optimisés si les images ont une taille en puissance de deux, nous avons fixé deux tailles cibles pour les imagerie à savoir 32×32 et 64×64 . Par conséquent, la fenêtre de segmentation a été partitionnée en 8×8 cases de dimensions 16×16 pixels comme le montre la figure 3.7.

3.3.1.3 Extraction d'imagerie

A partir de l'image carrée segmentée et de son masque, nous allons extraire un maximum d'imagerie avec un pas de seize pixels (une case de l'IHM *Mosaic*). Comme évoquée dans la section précédente, deux tailles d'imagerie sont extraites : 32×32 (soit quatre cases) et 64×64 (soit seize cases) pixels. Pour qu'une imagerie soit étiquetée « Thrombus », elle doit contenir au moins 90 % de pixels étiquetés « Thrombus » (tolérance d'une case sur seize). Si une image est composée d'au moins 5 % (soit un peu moins d'une case) de thrombus et de plus de 10 % de pixels étiquetés « Mixtes », alors l'imagerie est extraite et étiquetée « Mixtes ». Enfin, si l'ensemble des pixels de l'imagerie

sont étiquetés « A considérer » (*i.e.* au moins un clic), l'image est extraite et étiquetée comme « Tissu ». Ces pourcentages ont été déterminés de manière empirique et nécessiteraient sûrement une optimisation en cas de résultats positifs sur la classification des thrombus (procédure de validation croisée). La figure 3.8 affiche les neuf imagettes extraites de l'image de la figure 3.7.

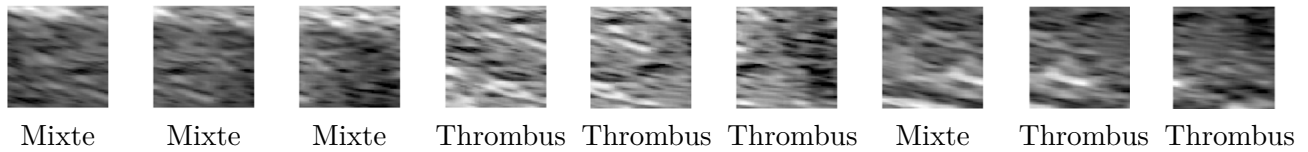


FIGURE 3.8 – Imagettes de taille 64×64 pixels (soit seize cases) et extraites à partir de l'image segmentée illustrée en figure 3.7 : 5 images contenant uniquement du thrombus et 4 images mixtes (thrombus, parois et tissu extérieur).

3.3.2 Echographies des glandes salivaires

Nous allons nous intéresser maintenant à l'extraction des imagettes à partir des échographies de glandes salivaires. Avant l'extraction, ces échographies doivent être segmentées pour déterminer la nature de l'imagette extraite. Comme pour les thrombi, les imagettes seront extraites automatiquement à partir de l'image segmentée et étiquetées selon trois catégories : glandes, mixtes et référence. Contrairement aux images de thrombi, la segmentation des glandes salivaires n'a pas été réalisée en créant une mosaïque de l'échographie mais en dessinant directement le contour de la glande sur l'échographie. Les glandes sont en effet plus volumineuses et de forme plus complexe que les thrombi. Ces segmentations ont été effectuées sous la tutelle du service de rhumatologie du CHRU de Brest. Des exemples de glandes segmentées sont affichées sur les figures 1.16 et 3.9.a.

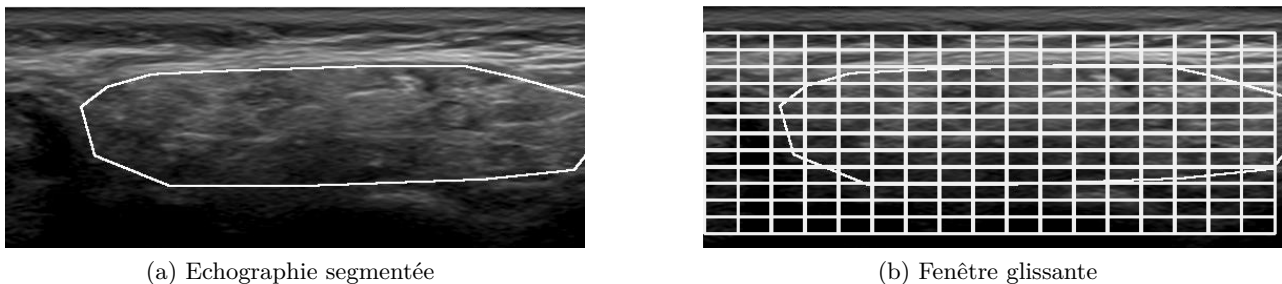


FIGURE 3.9 – Segmentation et extraction d'imagettes sur les échographies de glandes salivaires : la fenêtre glissante est extraite que si elle contient, soit 95 % de pixels de glande (étiquetée « Glande »), soit entre 40 et 55 % de glande (étiquetée « Mixte »).

Une fois l'échographie segmentée, les imagettes étiquetées « Glande » et « Mixte » sont extraites automatiquement de l'image. Le but étant la validation de nos approches dans l'optique de la caractérisation de la structure des thrombus, nous avons choisi les deux mêmes tailles cibles (32×32 et 64×64 pixels). Une fenêtre glissante parcourt l'échographie et extrait l'imagette si la fenêtre contient un certain pourcentage de pixels de glande : 95 % pour être « Glande », entre 40 et 55 % pour « Mixte ». Les 5 % de tolérance s'expliquent par le fait que le contour est dessiné sur le bord interne de la glande : quelques pixels de glandes peuvent donc se situer à l'extérieur du masque de segmentation (d'où également 40 % de pixels de glandes au lieu de 45 % qui pourrait paraître plus logique au vue de la borne maximale de 55 %). Enfin, il reste deux cas à traiter : le cas où aucune image n'est extraite et l'extraction d'une imagette de référence de chaque échographie. Ces deux cas sont traités manuellement avec la même IHM *Optio* qui est présenté dans la section suivante pour extraire les imagettes d'élastographies.

3.3.3 Elastographies des thrombi et des glandes salivaires

Pour les élastographies, nous avons tout d'abord pensé à extraire automatiquement les imagerie en s'aidant des ellipses dessinées par les praticiens. Cette procédure automatique est présentée en annexe B.3. Toutefois, en automatique, nous nous heurtons à trois problèmes : les zones sans données (trous dans l'élastographie), les acquisitions où l'image sans ellipse n'est pas disponible et le choix de l'imagerie de référence. En ce qui concerne les zones sans données, il faudrait donc les détecter et extraire une imagerie à un autre endroit où il y a du thrombus mais pas de trous. Pour pallier ces trois problèmes, l'IHM *Optio* a été développée pour extraire manuellement jusqu'à trois imagerie de la région d'intérêt (thrombus ou glande) et une imagerie de référence. Cette IHM, illustrée en figure 3.10, permet à son utilisateur de définir avec précision la position des imagerie qui seront extraites. Dans le cas où il n'est pas possible de placer le rectangle de sélection avec la taille désirée (32×32 ou 64×64 pixels), l'utilisateur peut alors choisir de réduire la dimension du rectangle (à cause de trous ou de la présence de l'ellipse). L'imagerie extraite de ce rectangle sera ensuite agrandie par effet miroir (*e.g.* un vecteur $[1 \ 2 \ 3]$ agrandie de deux valeurs par sa gauche donnera $[2 \ 1 \ 1 \ 2 \ 3]$) à la taille cible souhaitée. L'utilisateur peut ensuite décider s'il souhaite conserver ou non l'image agrandie.

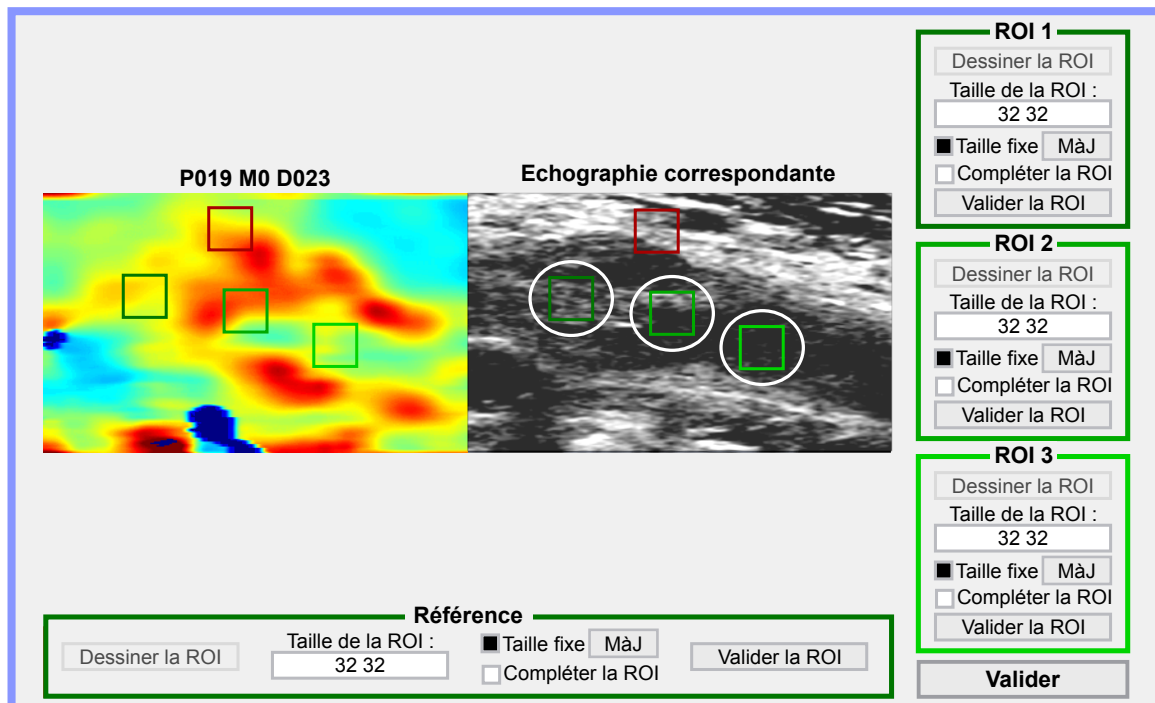


FIGURE 3.10 – IHM *Optio* pour extraire manuellement des imagerie : l'utilisateur peut extraire une imagerie de référence et jusqu'à trois imagerie de la région d'intérêt. La taille des imagerie extraites est fixée en entrée mais peut être modifiée suivant les besoins de l'acquisition (en décochant la case « Sized fixed »).

3.3.4 Conclusion sur la segmentation et l'extraction des imagerie

Cette section a décrit deux étapes clés de la préparation des données car elles vont conditionner les autres étapes de la chaîne de traitement à savoir l'extraction de descripteur et la classification. En effet, l'objectif est d'obtenir, en sortie des deux étapes présentées ici, des imagerie les plus pertinentes possible pour caractériser les thrombi (ou les glandes salivaires). Pour ce faire, il faut d'abord définir la notion de région d'intérêt. Dans la première partie de la thèse, nous ne considérons que le thrombus seul (intérieur de l'ellipse). A présent, nous avons décidé d'être un peu plus souple (tolérer les parois et un peu de tissu extérieur) afin de modifier la taille des images extraites et donc l'échelle des macrostructures du thrombus. Ensuite, nous avons ajouté l'extraction d'une image de référence pour chaque acquisition et qui sera utilisée par la suite pour normaliser les imagerie et/ou ses descripteurs.

Pour cela, nous avons développé deux stratégies d'extraction d'imagette en fonction du type d'image : semi-automatique pour les échographies et manuelle pour les élastographies. Les échographies sont en effet segmentées manuellement mais l'extraction est automatique.

3.4 Conclusion sur la préparation des bases de données

Ce chapitre a présenté l'ensemble des étapes qui permettent de préparer les images de thrombus pour les simulations décrites au chapitre 5. Nous avons pu constater que certains maillons de la chaîne de traitement ont évolué au cours de la thèse afin de mieux répondre à la problématique de la thèse (changements de protocole induits par la mise à jour du système, inclusion des parois des veines dans les imagettes et extraction d'imagettes de référence). Nous avons vu qu'une acquisition est composée de plusieurs captures d'écran, peut être transversale ou longitudinale, possède des métadonnées et des informations indiquées uniquement sur la capture elle-même. L'IHM *Lectio* a donc été développée pour rassembler les images d'une même acquisition et récupérer l'ensemble des données relatives à l'acquisition (paramètres, informations sur les patients et l'examen). Les échographies voient leurs histogrammes être égalisés pour tenter de limiter l'influence du gain qui peut varier d'un patient à l'autre. Puis, elles sont segmentées pour distinguer les régions contenant uniquement du thrombus des régions mixtes (avec l'IHM *Mosaic*) et pour pouvoir étiqueter les images qui sont extraites ultérieurement. Les élastographies, quant à elles, sont converties en carte de vitesse, et donc de dureté (ces deux grandeurs étant liées, cf. section 1.2.3.2). L'extraction des imagettes se fait de manière manuelle à l'aide de l'IHM *Optio* afin d'éviter les zones sans données et les annotations. Après avoir parcouru l'ensemble des acquisitions des bases de données (thrombi et glandes), la prochaine étape de la chaîne de traitement (voir figure 2.2) consiste à calculer les descripteurs des images avec les deux approches décrites dans le chapitre 4 et de les classer avec les algorithmes de la section 2.4.

Chapitre 4

Caractérisation des thrombi par approches multirésolutions et statistiques

Glossaire du chapitre 4

2D Deux Dimensions

ACP Analyse en Composantes Principales (ou *Principal Component Analysis* en anglais, PCA)

AME Apprentissage Multi-Echelle (ou multirésolution)

AMR Analyse MultiRésolution (ou multi-échelle)

ROI Région d'intérêt (*Region Of Interest*)

SO *Scattering Operator* : réseau convolutif de transformée en ondelettes et d'opérateurs contractants

SOS Statistiques d'Ordre Supérieur (*High Order Statistics* en anglais, HOS)

TFD Transformée de Fourier Discrète

Introduction

Au chapitre 2, nous avons pu constater que l'étude des images de thrombi en fonction de l'épidémiologie de la maladie est une approche innovante dans ce domaine. Par conséquent, l'état de l'art a été élargi à d'autres domaines d'applications biomédicales comme l'imagerie des glandes mammaires, prostatiques ou salivaires. Ce chapitre fournit également une liste non exhaustive de méthodes pour chaque étape de la chaîne de traitement (voir figure 2.2) dans le cadre d'une problématique de classification d'images. Dans ce chapitre, nous allons nous concentrer sur les deux approches que j'ai développées pour extraire des descripteurs des images de thrombi veineux. La première approche, décrite en section 4.1 exploite la théorie des ondelettes pour transformer l'espace de représentation des images : le *scattering operator*. La seconde approche, détaillée en section 4.2, exploite les statistiques d'ordre supérieur pour extraire de l'information pertinente des images. Enfin, dans la section 4.3, nous nous interrogerons sur la combinaison de ces deux approches de nature totalement différente pour améliorer le pouvoir discriminant de nos descripteurs en vue de la classification.

4.1 Approche multirésolution : le *Scattering Operator*

Une image acoustique (échographie ou élastographie) comporte souvent différentes textures qui dépendent du type de tissu observé (organes, vaisseaux, tendons, *etc.*) et de son état de santé (sain, bénin, malin, fibrosé, *etc.*). Ces images souffrent souvent de petites déformations locales dues aux caractéristiques du tissu et au procédé d'acquisition (approximations faites lors de la reconstruction de l'image). L'idée générale du *Scattering Operator* (SO) est de réduire la variabilité entre les images d'une même classe (même type et état de tissu) tout en conservant la variabilité entre les classes. Pour ce faire, il transforme l'espace de représentation des images pour que celui-ci devienne invariant aux petites déformations et aux translations. La section 4.1.1 revient sur le principe du SO puis la section 4.1.2 présente les opérateurs (dont les ondelettes) nécessaires pour construire le réseau convolutif du SO. La section 4.1.3, illustre deux manières de représenter les coefficients du SO. Enfin, nous verrons comment conférer au SO des propriétés d'invariance en rotation et en échelle. L'ensemble de ces sections s'appuie sur l'annexe C qui pose les bases mathématiques de l'Analyse MultiRésolution (AMR), sur les articles [22, 151, 161] et sur la thèse de Bruna [21].

4.1.1 Principe

Introduit par Mallat [22], le développement du SO vise particulièrement l'amélioration des techniques de classification automatique de signaux (par exemple les images). Par exemple, si l'on part d'un nuage de points représentant de deux classes d'images différentes mais non dissociables, l'objectif est de trouver une fonction Φ qui va changer le mode de représentation des points de telle sorte que les deux classes soient distinguables (voir figure 4.1).

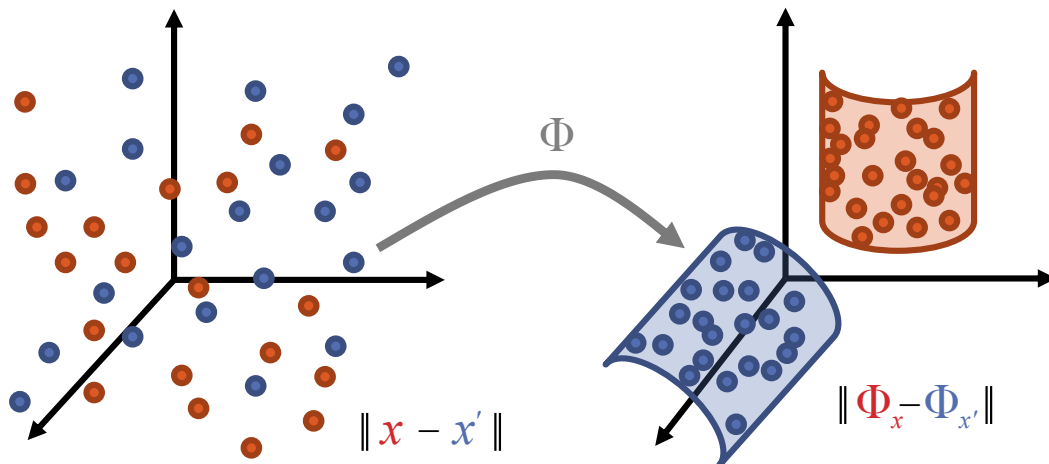


FIGURE 4.1 – Contraction de l'espace par l'opérateur Φ pour obtenir une meilleure représentation

En traitement d'image, une difficulté majeure des techniques de classification réside dans la possible grande variabilité à l'intérieur des classes d'image. Une part importante de cette variabilité est due à des transformations (translation, rotation, et changement d'échelle) rigides (identiques pour tous les pixels de l'image) des images de chaque classe. Cette variabilité est souvent non-informative dans le sens où elle ne caractérise pas un changement de classe. Pour obtenir une classification optimale, cette variabilité doit être maîtrisée, c'est-à-dire qu'une image x et la même image ayant subi une transformation rigide x' devraient être très proches dans l'espace de représentation de ces images pour la classification. Cela s'exprime mathématiquement en terme de distance entre deux images x et x' (ayant subi la transformation rigide) par :

$$\|\Phi_x - \Phi_{x'}\| = 0 \quad (4.1)$$

où Φ_x est l'image de x dans le nouvel espace de représentation et $\|\cdot\|$ est la norme euclidienne.

Les déformations non-rigides sont également une importante source de variabilité à l'intérieur des classes. Ces déformations peuvent mener à la classification de l'image dans une autre classe. Dans le cadre de l'algorithme de Mallat, nous nous limitons à des transformations non-rigides à modifications restreintes. En termes mathématiques, l'algorithme de Mallat utilise la propriété de continuité de Lipschitz relative aux déformations :

$$\exists C > 0, \forall \tau \forall x : \|\Phi_{x_\tau} - \Phi_x\| \leq C\|x\| \sup_u |\nabla\tau(u)| \quad (4.2)$$

avec :

- τ est un champ de déplacements (déformations) non constant des pixels de l'image x ;
- x_τ est l'image x transformée par τ ;
- $\nabla\tau$ est le tenseur du gradient de la déformation dont la norme $|\nabla\tau|$ mesure l'amplitude de la déformation et $\sup_u |\nabla\tau(u)|$ est l'amplitude de la déformation globale de l'image.

Cette propriété exprime simplement que pour une petite déformation de l'image, la différence des deux représentations est bornée par l'amplitude de la déformation (voir figure 4.1).

4.1.2 Ondelettes et propriétés du *Scattering Operator*

Le SO convolue chaque image d'entrée avec des ondelettes dilatées et orientées. Chaque ondelette est obtenue par rotation θ et dilatation j (appelée échelle) d'une ondelette mère ψ :

$$\psi_\lambda(u) = 2^{-2j}\psi(2^{-j}r_\theta^{-1}u) \quad (4.3)$$

où

- $\lambda = 2^j r_\theta$;
- $u = (u_1, u_2)^T$ est le vecteur des positions spatiales ;
- $r_\theta = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$ est la matrice de rotation d'angle θ (orientations)

Dans nos futures simulations décrites au chapitre 5, l'ondelette mère ψ est une ondelette de Morlet [22] qui n'est autre qu'une exponentielle complexe multipliée par une fenêtre gaussienne (*cf.* équation (C.8) en annexe C.1.3).

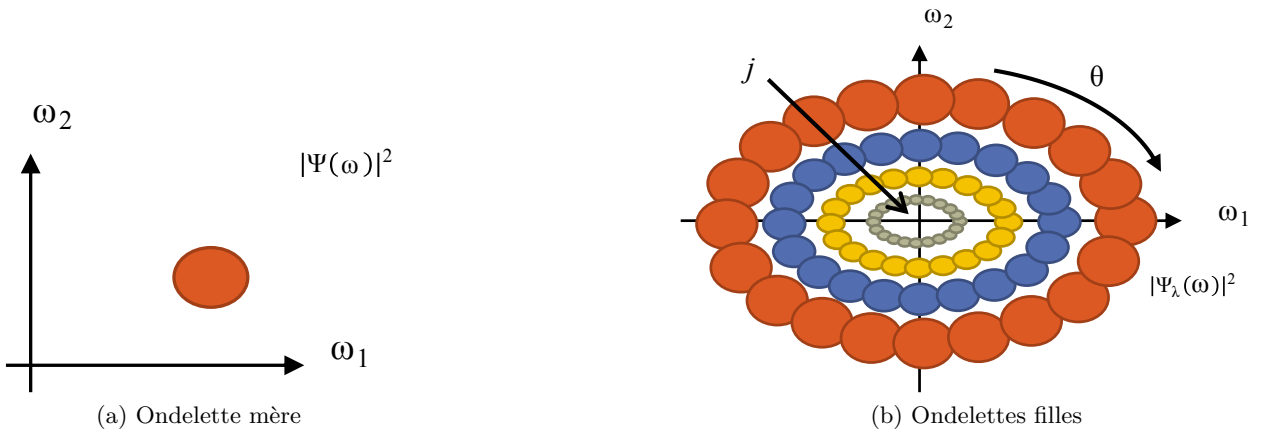


FIGURE 4.2 – Support fréquentiel d'une ondelette mère et balayage du plan fréquentiel avec des ondelettes directionnelles (dilatation et rotation) appelées ondelettes filles.

Dans le cas de la transformée d'images discrètes et réelles, les orientations θ sont discrétisées et seules les rotations positives r_θ sont considérées; on pose $\theta = l\pi/L$ avec $0 \leq l \leq L$. De la même manière l'échelle 2^j prend un nombre limité de valeurs $0 \leq j \leq J$. Les différentes ondelettes filles ψ_λ sont obtenues par dilatations proportionnelles à 2; on parle de dilatation dyadique. La figure 4.2 illustre le support fréquentiel (Ψ_λ est la transformée de Fourier de ψ_λ) d'une ondelette ψ_λ . Chacune de ces fonctions définit un filtre passe-bande dont la surface et la localisation fréquentielle dépendent de l'échelle j et de l'angle d'orientation θ . En faisant varier ces deux paramètres, l'ensemble de ces filtres constitue un banc de filtres dont la couverture mutuelle du plan fréquentiel permet de réaliser une Analyse MultiRésolution (AMR) de l'image.

Toutefois, la figure 4.2.b montre que l'information basse-fréquence n'est pas analysée par le banc de filtres passe-bande ψ_λ . Pour étudier le plan fréquentiel dans sa totalité, nous allons ajouter un filtre passe-bas ϕ_J , appelé ondelette père, à ce banc de filtres. Le filtre ϕ_J est défini par :

$$\phi_J(u) = 2^{-2J}\phi(2^{-J}u) \quad (4.4)$$

où u est le vecteur des positions spatiales et J est le niveau d'échelle maximal des filtres passe-bande. Dans nos simulations, ϕ est un filtre gaussien invariant en rotation [22] et son contenu fréquentiel est donc centré sur la fréquence nulle.

Au final, la transformée en ondelettes d'une image discrète x à l'échelle 2^J consiste à lui appliquer le banc de filtres défini par :

$$\{x * \phi_J, (x * \psi_\lambda, \lambda \in \Lambda_J)\} \quad (4.5)$$

avec $\Lambda_J = \{\lambda = 2^j r_\theta, 0 \leq j \leq J, \theta = \frac{l\pi}{L} \text{ avec } 0 \leq l \leq L\}$.

Les ondelettes ont la particularité d'être plus proches des signaux naturels que les sinusoides utilisées dans le domaine de Fourier et d'être stables aux petites déformations (une ondelette légèrement déformée ressemble toujours à une ondelette). Les ondelettes ne sont pas invariantes en translation mais cette invariance peut être obtenue en appliquant le module et un filtre passe-bas (l'ondelette père). Le module va déplacer l'information haute fréquence extraite par les ondelettes vers les basses fréquences puis le filtre passe-bas va nous donner l'information haute fréquence extraite invariante en translation.

L'invariance en translation, obtenue par l'usage de l'ondelette père ϕ_J , concerne des échelles plus petites que 2^J (où J est le niveau d'échelle). La variabilité spatiale est conservée pour des échelles plus grandes que 2^J . Le premier coefficient du SO de l'image x est défini de la façon suivante :

$$S[\emptyset]x = x * \phi_J \quad (4.6)$$

où $*$ désigne le produit de convolution et \emptyset signifie que le coefficient $S[\emptyset]x$, obtenu par filtrage passe-bas, n'a pas été filtré par un passe-bande (*i.e.* une ondelette dilatée et orientée).

$S[\emptyset]x$, appelé coefficient d'approximation est bien invariant par translation mais ne contient que l'information basse-fréquence. Toutefois, l'information perdue va être récupérée dans un autre temps d'analyse, en calculant les coefficients d'ondelettes $U[\lambda_1]x$ de l'image x définis pour chaque orientation θ et échelle j par :

$$U[\lambda_1]x = |x * \psi_{\lambda_1}| \quad (4.7)$$

avec $\lambda_1 = 2^{j_1} r_{\theta_1} \in \Lambda_J = \{\lambda = 2^j r_\theta, 0 \leq j \leq J, \theta = \frac{l\pi}{L} \text{ avec } 0 \leq l \leq L\}$.

Les coefficients $U[\lambda_1]x$ sont, par les propriétés des ondelettes, stables aux petites déformations mais ne sont pas invariants en translation. Cette invariance est obtenue en appliquant le filtre moyennneur ϕ_J à chaque transformée $U[\lambda_1]x$:

$$S[\lambda_1]x = U[\lambda_1]x * \phi_J = |x * \psi_{\lambda_1}| * \phi_J. \quad (4.8)$$

La famille d'invariants $S[\lambda_1]x$ ainsi formée constitue les coefficients du SO d'ordre un (premier niveau de décomposition).

L'utilisation du filtre passe-bas a engendré une perte d'information qu'on peut récupérer en appliquant à nouveau une transformée en ondelettes à chaque coefficient $U[\lambda_1]x$. Ce deuxième niveau de décomposition permet ainsi d'obtenir les coefficients d'ordre 2. La généralisation de cette technique définit une famille d'invariants en translation. L'ordre m de décomposition est calculé en suivant un chemin $p = (\lambda_1, \lambda_2, \dots, \lambda_m)$ et est défini par :

$$\begin{aligned} U[p]x &= U[\lambda_m] \dots U[\lambda_2] U[\lambda_1]x \\ &= |||x * \psi_{\lambda_1} | * \psi_{\lambda_2} | \dots * \psi_{\lambda_m} | \end{aligned} \quad (4.9)$$

La transformée $S[p]x$ à l'ordre m de l'image x est une image définie par :

$$\begin{aligned} S[p]x &= U[p]x * \phi_J \\ &= |||x * \psi_{\lambda_1} | * \psi_{\lambda_2} | \dots * \psi_{\lambda_m} | * \phi_J \end{aligned} \quad (4.10)$$

En pratique, une étape d'agrégation (ou *pooling* en anglais, cf. section 2.4.1.3) est appliquée à chaque transformée $S[p]x$ pour la réduire à une seule valeur (e.g. la fonction moyenne). Finalement, la sortie du SO avec deux ordres de décomposition est un vecteur défini par :

$$Sx = \left[\tilde{S}[\emptyset]x, \tilde{S}[\lambda_1]x, \tilde{S}[\lambda_1, \lambda_2]x \right] \quad (4.11)$$

où chaque coefficient $\tilde{S}[p]x$ est le résultat de l'agrégation de l'image $S[p]x$ obtenue pour un chemin p donné. La taille du vecteur Sx dépend donc du nombre d'ordres, d'échelles et d'orientations.

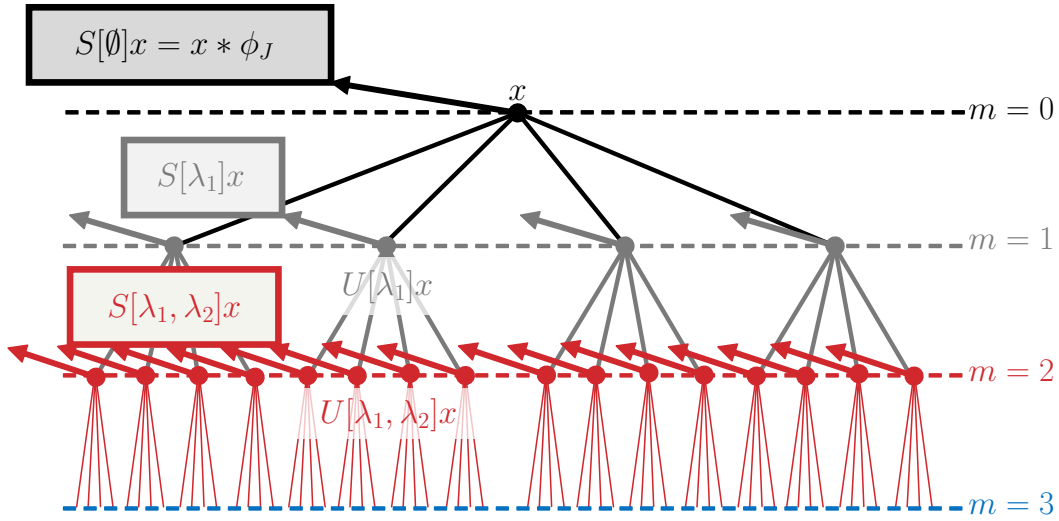


FIGURE 4.3 – Représentation du SO en réseau de convolutions pour $m = 3$ niveaux de décompositions et deux échelles et deux orientations.

La transformée du SO peut être vue comme un réseau de (filtres de) convolutions qui utilise les ondelettes comme filtres. La figure 4.3 illustre cette structure de réseau pour deux échelles et deux orientations. Il est à noter que la structure serait identique si on réalisait une décomposition basée sur quatre échelles et une orientation, ou quatre orientations et une échelle (chaque nœud de la structure se décompose en quatre branches). L'ensemble des coefficients obtenus aux différents ordres de décomposition forme une représentation stable aux déformations, invariante en translation et donc efficace en vue d'une classification. Elle préserve également l'énergie présente dans l'image à analyser. En effet, on peut montrer que l'énergie des scattering coefficients décroît de façon exponentielle en fonction de l'ordre de décomposition m et que quatre ordres de décomposition suffisent pour récupérer plus de 99 % de l'énergie de l'image [22]. L'utilisation du logarithme avant l'agrégation spatiale permet de linéariser cette décroissance d'énergie et de renforcer l'information présente dans les ordres supérieurs.

4.1.3 Représentation des transformées du *scattering operator*

Il est possible de représenter les coefficients $S[p]x$ dans le plan des fréquences spatiales normalisées $(\omega_1, \omega_2) = \frac{2\pi}{F_E}(f_1, f_2)$; F_E est la fréquence d'échantillonnage), sous forme de disque [22]. Le disque est découpé en secteurs représentatifs d'une zone fréquentielle. A chacun de ces secteurs, on peut associer un coefficient du SO. La figure 4.4 illustre la division du plan fréquentiel en quadrants.



FIGURE 4.4 – Représentation des coefficients du scattering dans le plan fréquentiel.

La figure 4.5 montre un exemple pour une image composée de rayures orientées à 45° . A l'ordre un, la valeur des coefficients est maximale pour l'orientation des rayures (45° et 135°) et pour l'échelle la plus grande ($j = 2$). A l'ordre deux, le maximum des coefficients est obtenu dans la direction 45° et 135° . La discrétisation plus fine permet une meilleure sélectivité fréquentielle. On peut également représenter les coefficients du SO comme sur la figure 4.6.

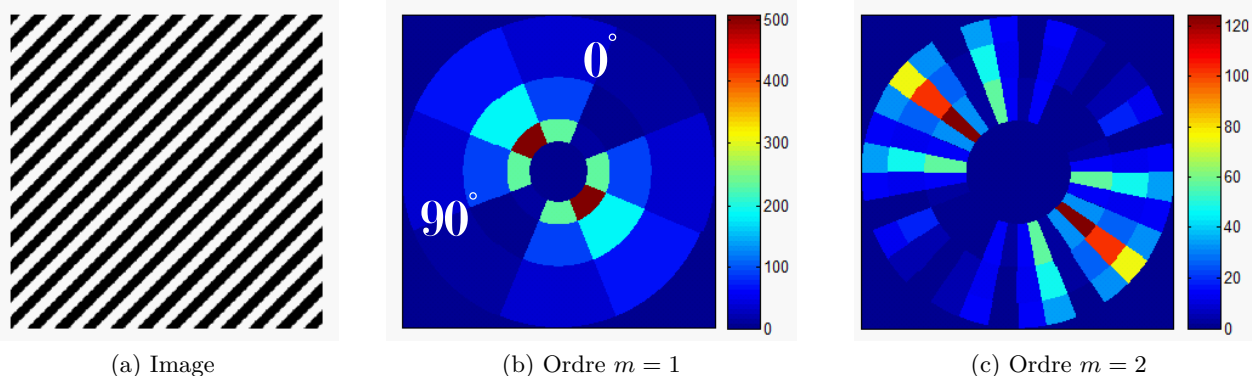


FIGURE 4.5 – Représentation des coefficients dans le disque fréquentiel aux ordres $m = 1$ et $m = 2$ pour une image composée de rayures orientées à 45° .

Les trois colonnes correspondent chacune à un ordre de décomposition $m = 0, 1, 2$. En première ligne, est représentée la valeur des coefficients $S[p]x$ en fonction du numéro de coefficients, c'est-à-dire pour chaque séquence p d'échelles et d'orientations possibles. La seconde permet de connaître les échelles correspondant à p . A l'ordre zéro, seul un filtre passe-bas à l'échelle indiquée dans la première colonne est appliqué. A l'ordre un, on peut voir les échelles utilisées pour la décomposition en ondelettes ainsi que l'échelle du filtre passe-bas appliqué *a posteriori*. A l'ordre deux, on effectue une nouvelle transformée en ondelettes sur chacune des sorties de la transformée en ondelettes de l'ordre précédent. On peut ainsi lire l'échelle de la première transformée en bleu et celle de la seconde en vert. Enfin, la troisième ligne représente les orientations utilisées dans les transformées en ondelettes.

Pour l'exemple représenté en figure 4.6, les valeurs maximales à l'ordre un sont localisées pour l'orientation 45° et pour l'échelle la plus grande ($j = 2$). A l'ordre deux, deux filtres sont passés sur l'image. Concernant les scattering coefficient, on observe plusieurs maxima. Ces derniers sont tous localisés pour une orientation de 45° d'un des deux filtres. La valeur maximale est due à la conjonction des deux filtres orientés à 45° et de l'échelle maximale.

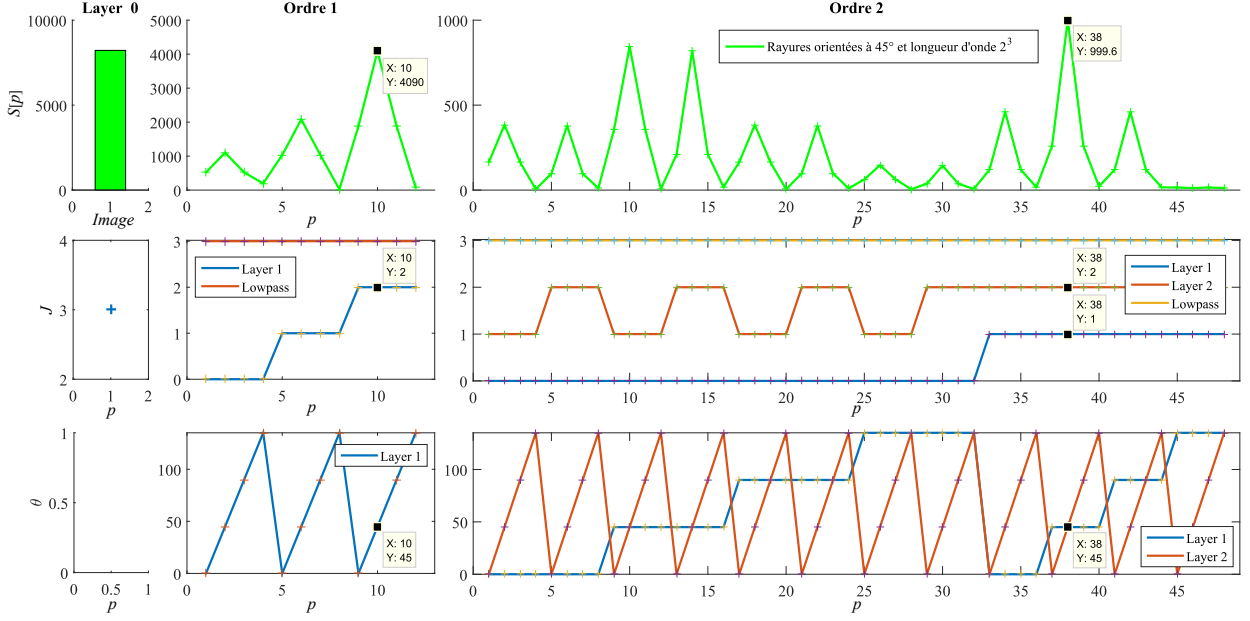


FIGURE 4.6 – Représentation des coefficients du SO en fonction de l'ordre de décomposition $m = 0, 1, 2$ (haut) ; Facteur d'échelle j utilisé (milieu) ; Orientation θ utilisée (bas).

4.1.4 Invariances en rotation et en échelle

La section précédente a montré comment le SO permettait de construire une nouvelle représentation de l'image de dimension inférieure, stable aux petites déformations et invariante en translation. Les auteurs de [151] proposent d'aller plus loin en rendant cette représentation invariante en rotation et en échelle.

4.1.4.1 Invariance en rotation :

Pour construire cette invariance en rotation, le *scattering operator* calcule d'abord les coefficients $S[\emptyset]x$ et $U[\lambda_1]x$ tels qu'ils sont définis par les équations (4.6) et (4.7). Ensuite, en plus du filtre moyenneur spatial ϕ_J , pour chaque échelle j_1 fixée, les coefficients obtenus pour chaque rotation θ_1 sont moyennés. Pour indiquer la dépendance en θ_1 , $U[\lambda_1]x$ est noté $U[j_1, \theta_1]x$ dans cette section. Les coefficients d'ordre un du SO avec l'invariance en rotation, notés $S[j_1]$ (pour indiquer qu'ils ne dépendent plus de θ_1), sont définis par :

$$S[j_1] = U[j_1, \theta_1]x \otimes \Phi_J \quad (4.12)$$

où :

- $\Phi_J(u, \theta) = \overline{\phi}(\theta)\phi_J(u)$ pour toute position spatiale u et pour toute position angulaire θ ;
- $\overline{\phi} = (2\pi)^{-1}$ pour toute position angulaire θ ;
- \otimes indique que les convolutions sont réalisées suivant les orientations (positions angulaires).

En pratique, la convolution \otimes consiste à calculer les coefficients d'ondelettes pour plusieurs orientations à une échelle donnée (avec des filtres de convolution) puis de prendre leur moyenne [151]. L'ondelette Φ_J permet donc de réaliser une moyenne « angulaire » (*i.e.* une moyenne suivant l'ensemble des orientations considérées) et une moyenne spatiale proportionnelle à 2^J . De cette façon, les coefficients $S[\lambda_1]x$ de l'ordre un sont invariants en translation et en rotation.

Pour les ordres supérieurs, des ondelettes tridimensionnelles sont appliquées pour récupérer l'information perdue par les filtres moyenneurs spatial et en rotation. Ces ondelettes sont définies par trois produits séparables. Les filtres spatiaux passe-bande ψ_{j_2, θ_2} ou passe-bas ϕ_J sont multipliés par des ondelettes complexes 2π -périodiques $\bar{\psi}_k$ [106] ou par $\bar{\phi}$:

$$\Psi_{\theta_2, j_2, k_2}(u, \theta) = \psi_{j_2, \theta_2}(u) \bar{\psi}_{k_2}(\theta) \quad (4.13)$$

$$\Psi_{0, J, k_2}(u, \theta) = \phi_J(u) \bar{\psi}_{k_2}(\theta) \quad (4.14)$$

$$\Psi_{\theta_2, j_2, 0}(u, \theta) = \psi_{j_2, \theta_2}(u) \bar{\phi}(\theta) \quad (4.15)$$

où k_2 représente l'échelle angulaire de l'ondelette 2π -périodique, j_2 l'échelle de l'ondelette spatiale, θ_2 l'orientation de l'ondelette spatiale, u les positions spatiales et θ les positions angulaires.

La famille d'invariants $U[j_2, \theta_2, k_2]x$ est donc définie par :

$$U[j_2, \theta_2, k_2]x = |U[j_1, \theta_1]x \otimes \Psi_{\theta_2, j_2, k_2}| \quad (4.16)$$

Ensuite, comme pour l'ordre un, les coefficients $U[j_2, \theta_2, k_2]x$ sont moyennés spatialement et angulairement par Φ_J :

$$S[j_2, \theta_2, k_2]x = U[j_2, \theta_2, k_2]x \otimes \Phi_J = |U[j_1, \theta_1]x \otimes \Psi_{\theta_2, j_2, k_2}| \otimes \Phi_J \quad (4.17)$$

Finalement, après l'agrégation spatiale, la sortie du SO avec deux ordres de décomposition est un vecteur de coefficients invariants en translation et rotation défini par :

$$Sx = \left[\tilde{S}[\emptyset]x, \tilde{S}[j_1, \theta_1]x, \tilde{S}[j_1, \theta_1, j_2, \theta_2, k_2]x \right] \quad (4.18)$$

La figure 4.7 montre les coefficients d'ordre un (avant l'agrégation spatiale) calculés sans et avec l'invariance en rotation. Le SO a été appliqué avec quatre échelles et quatre orientations. L'ajout de l'invariance en rotation permet bien d'obtenir tous les contours des objets (verticaux, horizontaux, diagonaux, *etc.*). Son intérêt dépend donc des types de textures que l'on souhaite classifier puisque l'orientation des motifs des textures n'est plus un critère discriminant.

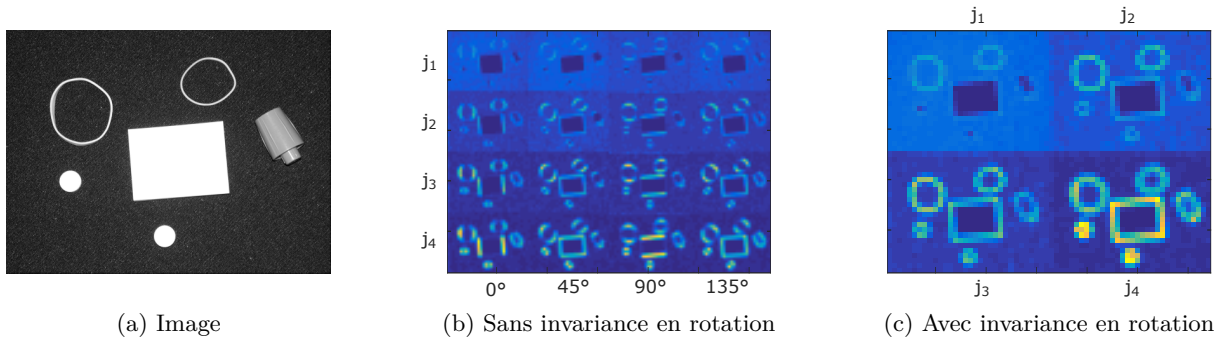


FIGURE 4.7 – Intérêt de l'invariance en rotation : chaque bloc représente les coefficients obtenus avec quatre échelles et, pour la figure (b), avec quatre orientations (0° , 45° , 90° et 135°).

4.1.4.2 Invariance en échelle :

Précédemment, nous avons construit une représentation de l'image invariante en rotation et en translation de taille inférieure à 2^J :

$$Sx = \left[\tilde{S}[\emptyset]x, \tilde{S}[j_1, \theta_1]x, \tilde{S}[j_1, \theta_1, j_2, \theta_2, k_2]x \right]$$

Ce vecteur Sx n'est pas invariant en échelle : deux images d'une même classe, mais avec une différence d'échelle supérieure à 2^J , peuvent donc être classées dans deux ensembles disjoints.

Pour ce faire, on va créer une nouvelle image x_i à partir de l'image d'origine x en modifiant l'échelle de l'image x suivant un pas de 2^i (*i.e.* un « zoom »). Pour rendre notre représentation invariante en échelle, on peut calculer Sx_i pour différentes échelles i et effectuer la moyenne de ces vecteurs. De plus, dans le cas d'une classification supervisée, il est possible d'augmenter la taille de la base d'apprentissage en considérant chaque image x_i indépendamment. On parle alors d'Apprentissage Multi-Echelle (AME) [151].

4.1.5 Conclusion sur le *scattering operator*

Dans cette section, nous avons découvert une première approche pour décrire nos images de thrombi de manière plus synthétique et en limitant la perte d'information : le *scattering operator*. Cette approche, qui s'appuie sur la théorie des ondelettes, analyse les images un peu à la manière de notre cortex visuel en étudiant d'abord grossièrement l'image puis en élevant progressivement le niveau de détail. Ce procédé rappelle celui de l'apprentissage profond évoqué en section 2.4.1.3 où les premières couches du réseau s'intéressent aux caractéristiques grossières (*e.g.* véhicule, animal, plante) et les couches suivantes à des niveaux de détails de plus en plus précis (*e.g.* voiture, camion, moto puis cabriolet, berline, compacte). Pour le SO, le niveau de détail va de pair avec la taille de l'échelle et l'ordre de décomposition. L'autre intérêt des ondelettes orientées combinées avec un module et un filtre passe-bas est que la transformation du SO assure une certaine stabilité face aux petites déformations et aux petites translations. Cette stabilité permet au SO de réduire la variabilité intra-classe et donc de réduire la dimension de l'espace tout en conservant le pouvoir discriminant des images. Cette approche fonctionne particulièrement bien pour distinguer des images texturées (matériaux [22], fonds marins [161]). Dans le cas où l'orientation et l'échelle des éléments de textures ne sont pas discriminantes, il est également possible de rendre la représentation du SO invariante à ce type de transformations (voir section 4.1.4). Au sujet de la problématique de notre projet, la question est donc de savoir si la texture des images de thrombi veineux peut être corrélée avec la structure des thrombi et le contexte de leur formation (cause, embolie pulmonaire, âge). La section suivante présente une approche totalement différente qui s'intéresse aux statistiques des images.

4.2 Approches par statistiques

Le traitement du signal (monodimensionnel ou bidimensionnel comme les images) utilise la théorie des probabilités qui permet, en particulier, la construction de modèles aléatoires. Le comportement des signaux aléatoires peut être partiellement ou complètement décrit à travers des lois de probabilité. La plupart des techniques classiques de traitement du signal sont basées sur une description simplifiée des signaux faisant appel aux moments d'ordres un ou deux (moyenne, variance, écart-type, ou corrélation). Toutefois, le développement des moyens de traitement a permis de prendre en compte une description plus subtile des signaux en utilisant des statistiques d'ordre supérieur à deux. Ces statistiques, développées en premier lieu pour l'astronomie, la prospection sismique et les communications, permettent d'obtenir des informations qui seront utilisables pour résoudre certains problèmes non accessibles à l'ordre deux comme la séparation aveugle de sources [39, 88, 130]. De nos jours, on les retrouve dans de nombreux domaines d'application [115] comme la géophysique, la parole [62, 111], le sonar, le radar et plus récemment le traitement d'image [103, 147] ou le biomédical [8].

L'annexe D propose une introduction aux statistiques d'ordre supérieur et à leur application sur des signaux aléatoires dans le domaine temporel. Cet annexe définit notamment les notions de moments et de cumulants pour des variables aléatoires, des vecteurs aléatoires, des processus aléatoires et des signaux temporels. Dans cette section, nous proposons d'utiliser ces statistiques d'ordre supérieur pour décrire nos images de thrombi avant l'application d'outils de classification. Une première approche pourrait consister à vectoriser l'image et la considérer comme un signal à temps discret. Cette approche est courante en traitement d'images et elle est notamment appliquée avec succès pour classer automatiquement les chiffres manuscrits par des réseaux complètement connectés [90]. Toutefois, avec cette approche, nous perdrons la cohérence spatiale entre les lignes et les colonnes, ce qui pourrait s'avérer comme une perte d'information importante pour caractériser la structure des thrombi. Une seconde approche consiste à traiter l'image comme un signal à temps discret à deux dimensions et à étendre les définitions des moments et des cumulants temporels (multicorrélations) vues à la section D.2.

4.2.1 Moments et cumulants spatiaux

En considérant l'image I comme un signal à temps discret, son moment spatial d'ordre q est défini par :

$$\mathcal{M}_q[I](i, j) = \mathbb{E} [I(m, n)I(m + i_1, n + j_1) \dots I(m + i_{q-1}, n + j_{q-1})] \quad (4.19)$$

où $i = (i_1, i_2, \dots, i_{q-1})$, $j = (j_1, j_2, \dots, j_{q-1})$ sont les retards, m et n représentent les pixels de l'image.

De la même façon, le cumulants spatial d'ordre q est :

$$\mathcal{C}_q[I](i, j) = \mathcal{C} [I(m, n), I(m + i_1, n + j_1), \dots, I(m + i_{q-1}, n + j_{q-1})] \quad (4.20)$$

En pratique, les images que nous utiliserons seront centrées avant le calcul des multicorrélations afin de simplifier les calculs. En adaptant l'estimateur précédent au cas 2D, les moments spatiaux d'ordre q d'une image I de taille $M \times N$ seront approchés par :

$$\widehat{\mathcal{M}}_q[I](i, j) = \frac{1}{N \times M} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} I(m, n)I(m + i_1, n + j_1) \dots I(m + i_{q-1}, n + j_{q-1}) \quad (4.21)$$

avec $(i_k, j_k)_{1 \leq k < q} \in [1 - M, M - 1] \times [1 - N, N - 1]$. Les cumulants spatiaux peuvent ensuite être calculés à partir des moments estimés par la formule de Leonov et Shirayev (équation (D.30)).

4.2.2 Autocorrélation, bicorrélation et tricorrélation

Dans cette sous-section, nous allons détailler les multicorrélations d'ordres deux, trois et quatre qui sont appelées respectivement *autocorrélation*, *bicorrélation* et *tricorrélation*. L'objectif ici est d'illustrer les définitions (parfois complexes) que nous avons vues à la sous-section précédente et de mieux comprendre la nature des descripteurs statistiques que nous allons utiliser pour extraire de l'information des images de thrombi.

Ordre deux :

Par définition, l'autocorrélation (ou cumulante d'ordre deux) spatiale d'une image I de taille $M \times N$ est donnée par :

$$\widehat{\mathcal{C}}_2[I](i_1, j_1) = \widehat{\mathcal{M}}_2[I](i_1, j_1) = \frac{1}{N \times M} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} I(m, n)I(m + i_1, n + j_1) \quad (4.22)$$

qui est donc une image de taille $(2M - 1) \times (2N - 1)$. Un exemple d'autocorrélation d'une image est illustrée en figure 4.8.

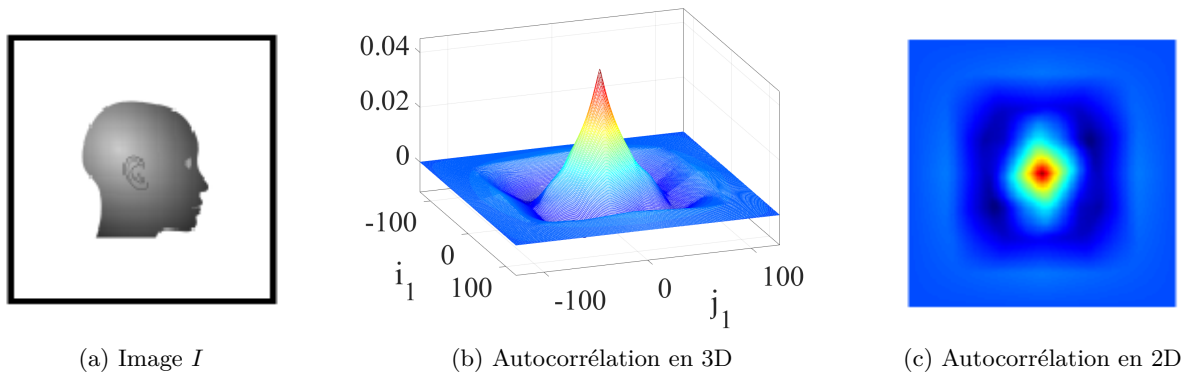


FIGURE 4.8 – Autocorrélation d'une image : l'image (c) correspond à la vue de dessus de la figure (b) (i_1 en profondeur, j_1 en horizontal et $\widehat{\mathcal{C}}_2[I](i_1, j_1)$ en hauteur).

La valeur centrale de l'image $\widehat{\mathcal{C}}_2[I](0, 0)$ n'est autre que la variance (ou l'écart-type si on prend sa racine) de l'image. Il est également à noter que, grâce aux propriétés d'ergodicité et de stationnarité (qui ne sont pas forcément vérifiées en pratique), il est possible d'estimer l'autocorrélation par la transformée de Fourier inverse de la densité spectrale de puissance (module au carré de la transformée de Fourier de l'image). L'astuce du passage par Fourier permet d'accélérer considérablement les temps de calcul pour calculer l'autocorrélation (cette astuce est également valable sur les signaux temporels).

Comme pour les couples de variables aléatoires et les signaux temporels, il est possible de comparer la ressemblance de deux images I et J de même taille $M \times N$ avec la corrélation croisée :

$$\widehat{\mathcal{C}}_2[I, J](i_1, j_1) = \widehat{\mathcal{M}}_2[I, J](i_1, j_1) = \frac{1}{N \times M} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} I(m, n)J(m + i_1, n + j_1) \quad (4.23)$$

La figure 4.9 montre comment la corrélation croisée, calculée en passant par le domaine de Fourier, pourrait être utilisée pour faire du recalage d'images.

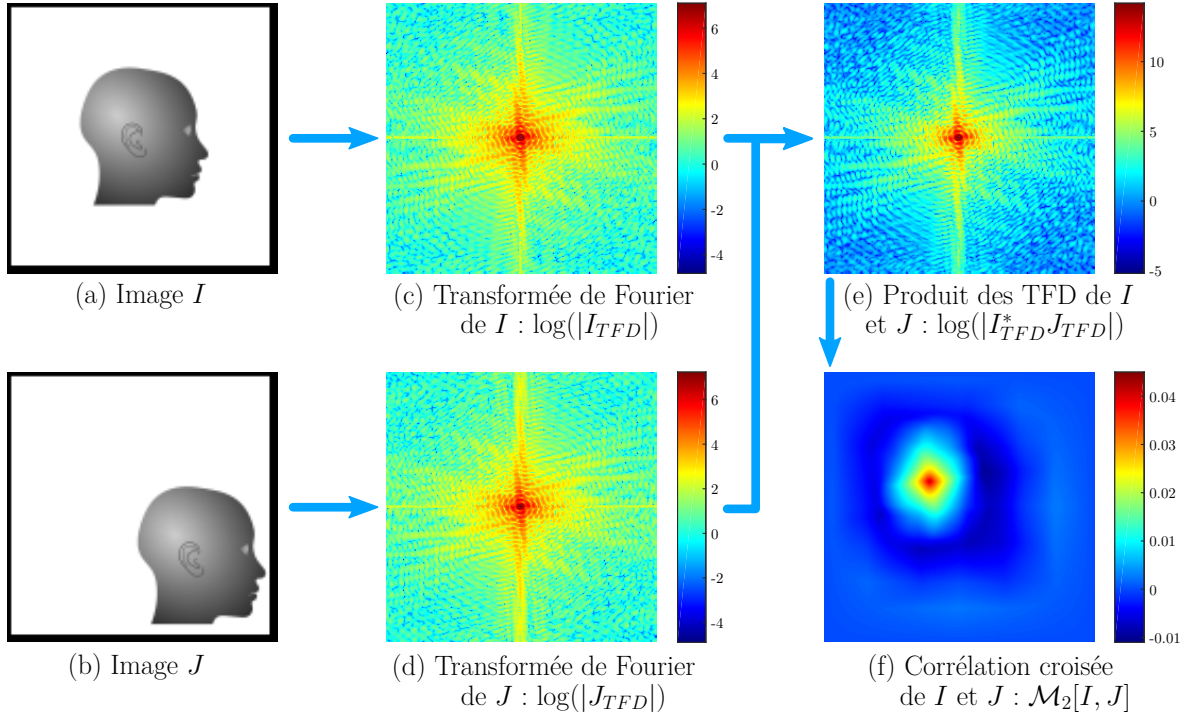


FIGURE 4.9 – Corrélation croisée de deux images I et J calculée par passage dans le domaine de Fourier : les images (c), (d) et (e) ne représentent pas directement les Transformées de Fourier Discrètes (TFD) des images et leur produit (si $J = I$, on aurait pu parler de densité spectrale de puissance de I) mais le logarithme de leur module pour obtenir un meilleur rendu.

Ordre Trois :

La bicorrélation de l'image I est :

$$\widehat{\mathcal{C}}_3[I](i_1, j_1, i_2, j_2) = \widehat{\mathcal{M}}_3[I](i_1, j_1, i_2, j_2) = \frac{1}{N \times M} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} I(m, n) I(m + i_1, n + j_1) I(m + i_2, n + j_2) \quad (4.24)$$

et est donc un tenseur d'ordre quatre composé de $(2M - 1)^2 \times (2N - 1)^2$ valeurs.

Le centre de ce tenseur d'ordre quatre, $\mathcal{C}_3[I](0, 0, 0, 0)$, divisé par l'écart-type σ_I au cube correspond au coefficient d'asymétrie (ou *skewness*) de l'image I :

$$\xi_I = \frac{\mathcal{C}_3[I](0, 0, 0, 0)}{\sigma_I^3} \quad (4.25)$$

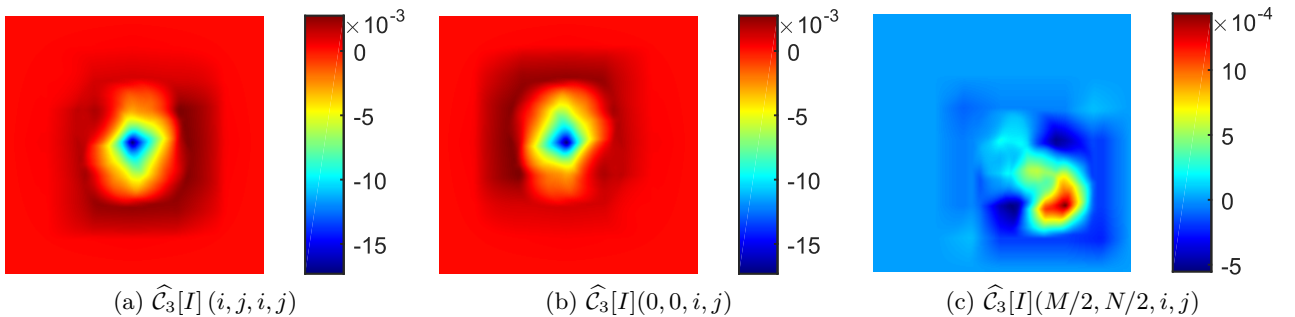


FIGURE 4.10 – Bicorrélation de l'image I de taille $M \times N$ illustrée à la figure 4.8.a suivant trois plans de coupe (i et j représentent les deux axes des plans de coupe de la bicorrélation).

La figure 4.10 représente différents plans de coupe de la bicorrélation de l'image de la figure 4.8.a. Comme pour l'autocorrélation, il est possible de calculer les bicorrélations croisées entre trois images I , J et K (ou de deux images en prenant $J = I$ ou $K = J$).

Ordre quatre :

La tricorrélation de l'image I est donnée par :

$$\widehat{\mathcal{C}}_4[I](i_1, j_1, i_2, j_2, i_3, j_3) = \widehat{\mathcal{M}}_4[I](i_1, j_1, i_2, j_2, i_3, j_3) \quad (4.26)$$

$$- \mathcal{M}_2[I](i_1, j_1) \mathcal{M}_2[I](i_2 - i_3, j_2 - j_3) \quad (4.27)$$

$$- \mathcal{M}_2[I](i_2, j_2) \mathcal{M}_2[I](i_1 - i_3, j_1 - j_3) \quad (4.28)$$

$$- \mathcal{M}_2[I](i_3, j_3) \mathcal{M}_2[I](i_1 - i_2, j_1 - j_2) \quad (4.29)$$

avec

$$\widehat{\mathcal{M}}_4[I](i_1, j_1, i_2, j_2, i_3, j_3) = \frac{1}{N \times M} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} I(m, n) I(m + i_1, n + j_1) I(m + i_2, n + j_2) I(m + i_3, n + j_3) \quad (4.30)$$

La tricorrélation est donc un tenseur d'ordre six composé de $(2M - 1)^3 \times (2N - 1)^3$ valeurs.

Le centre de ce tenseur d'ordre six, $\mathcal{C}_4[I](0, 0, 0, 0, 0, 0)$, divisé par l'écart-type σ_I à la puissance quatre correspond au coefficient d'aplatissement (ou *kurtosis*) de l'image I :

$$\kappa_I = \frac{\mathcal{C}_4[I](0, 0, 0, 0, 0, 0)}{\sigma_I^4} \quad (4.31)$$

La figure 4.11 représente différents plans de coupe de la tricorrélation de l'image de la figure 4.8.a.

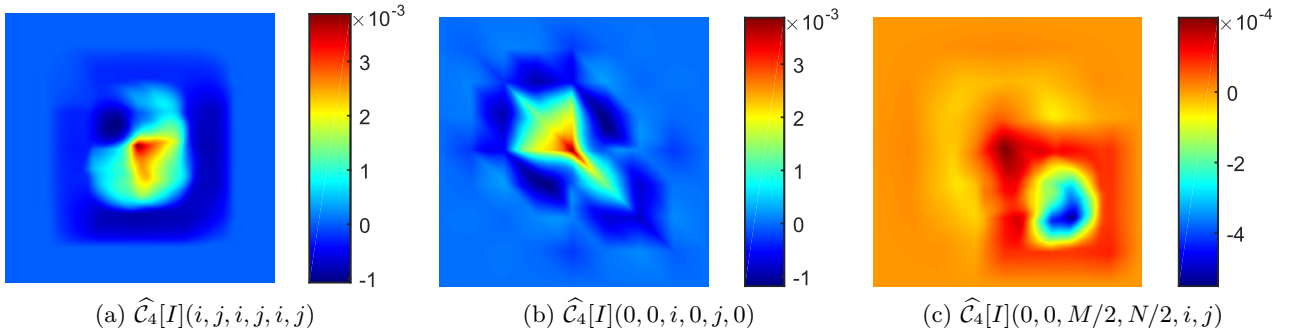


FIGURE 4.11 – Tricorrélation de l'image I de taille $M \times N$ illustrée à la figure 4.8a suivant trois plans de coupe (i et j représentent les deux axes des plans de coupe de la tricorrélation).

4.2.3 Construction des descripteurs

Comme nous avons pu le constater à la section précédente, plus la multicorrélation est d'ordre élevé, plus l'ordre du tenseur, et donc sa taille, est importante. A partir de l'ordre trois, l'espace mémoire peut notamment devenir saturé et/ou le temps de calcul exploser : pour une image de 64×64 pixels, sa bicorrélation comprendra plus de 260 millions de valeurs (*cf.* équation (4.24)). Par conséquent, à partir de l'ordre trois, il est compliqué, voire impossible, de calculer l'ensemble des valeurs. Parmi les solutions possibles, on pourrait sous-échantillonner les images ou réduire leur taille. Cependant, même en réduisant l'image à 16×16 pixels, la bicorrélation contiendrait encore presque un million de points. Une autre piste serait de calculer des valeurs seulement suivant certains axes privilégiés, à la manière de l'Analyse en Composantes Principales (ACP). Cardoso propose notamment une technique pour réduire

le tenseur d'ordre quatre des cumulants statistiques par une matrice générée à partir des vecteurs propres du tenseur (*Joint Approximation Diagonalization of Eigenmatrices* [25, 24]). Les auteurs de [48] présentent également un algorithme pour diagonaliser simultanément des tranches tensorielles d'ordre trois (*Simultaneous Thrid-Order Tensor Diagonalization*). De même, Comon propose de diagonaliser le tenseur d'ordre trois en maximisant la trace du tenseur [37, 38, 39]. Ces approches permettent de réaliser une *contraction de tenseur* (e.g. un passage d'un tenseur d'ordre trois à une matrice), qui, une fois adaptée aux multicorrélations, pourrait s'avérer intéressante.

Cependant, dans un premier temps, nous avons décidé d'opter pour une approche plus simple qui consiste à calculer les multicorrélations uniquement sur les deux dernières dimensions (i_{q-1}, j_{q-1}) du tenseur d'ordre q . Par ailleurs, de par la définition des multicorrélations, prendre (i_1, j_1) ou (i_2, j_2) donnerait exactement la même sortie. Avec, en entrée, des images de taille $M \times N$, nous obtiendrions donc, en sortie, les images de taille $(2M-1) \times (2N-1)$ suivantes : $\widehat{\mathcal{C}}_2[I](i, j)$, $\widehat{\mathcal{C}}_3[I](0, 0, i, j)$, $\widehat{\mathcal{C}}_4[I](0, 0, 0, i, j)$, etc. Pour chaque image, nous multiplierions donc quasiment par quatre le nombre de coefficients. En prévision de la classification, pour limiter les ressources mémoires et les temps de calculs, il est donc souhaitable de réduire encore davantage le nombre de coefficients.

Pour décrire les images de thrombi de manière plus synthétique, nous proposons de calculer les multicorrélations suivant un (ou plusieurs) axe(s) privilégié(s) sur le plan constitué des deux dernières dimensions du tenseur. Comme les imagerie de thrombi extraites sont de taille carrée ($N \times N$ pixels), les deux dernières dimensions du tenseur formeront donc une matrice carrée de taille $(2N-1) \times (2N-1)$. Le choix de ces axes s'est porté sur ceux qui paraissent les plus élémentaires et intuitifs : trois axes horizontaux, trois axes verticaux et les deux diagonales. Ces axes sont définis par la figure 4.12. Voici quelques exemples de vecteurs descripteurs que nous pouvons obtenir à partir de ces axes de multicorrélation :

- L'autocorrélation d'axe τ_1 est un vecteur de taille $2N-1$ obtenu en posant $(i_1, j_1) = (\tau, \tau)$ dans l'équation (4.22) :

$$\widehat{\mathcal{C}}_2[I](\tau_1) = \widehat{\mathcal{C}}_2[I](\tau, \tau) = \frac{1}{N \times N} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} I(m, n)I(m + \tau, n + \tau) \quad (4.32)$$

où τ est le vecteur $\tau = (1 - N, 2 - N, \dots, -1, 0, 1, \dots, N - 2, N - 1)^T$. Comme le montre la figure 4.12, le vecteur $\widehat{\mathcal{C}}_2[I](\tau_1)$ représente donc la diagonale de la matrice d'autocorrélation.

- L'autocorrélation axe τ_2 est :

$$\widehat{\mathcal{C}}_2[I](\tau_2) = \widehat{\mathcal{C}}_2[I](\tau', \tau') = \widehat{\mathcal{C}}_2[R_{90}(I)](\tau, \tau) \quad (4.33)$$

où τ' signifie qu'il faut calculer l'autocorrélation sur l'image $R_{90}(I)$, la rotation de 90 degré de l'image I . Le vecteur $\widehat{\mathcal{C}}_2[I](\tau_2)$ correspond donc à la diagonale opposée à $\widehat{\mathcal{C}}_2[I](\tau_1)$.

- La bicorrélation d'axe τ_3 est obtenue à partir de l'équation (4.24) :

$$\widehat{\mathcal{C}}_3[I](\tau_3) = \widehat{\mathcal{C}}_3[I](0, 0, 0, \tau) = \frac{1}{N \times N} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} I(m, n)I(m, n)I(m, n + \tau) \quad (4.34)$$

- La bicorrélation d'axe τ_6 est composée des vecteurs obtenus avec les axes τ_3 et τ_4 :

$$\begin{aligned} \widehat{\mathcal{C}}_3[I](\tau_6) &= \left(\widehat{\mathcal{C}}_3[I](\tau_3), \widehat{\mathcal{C}}_3[I](\tau_4) \right)^T \\ &= \left(\widehat{\mathcal{C}}_3[I](0, 0, 0, \tau), \widehat{\mathcal{C}}_3[I](0, 0, \tau, 0) \right)^T \end{aligned} \quad (4.35)$$

avec τ variant de $1 - N$ à $N - 1$, soit un vecteur de taille $2(2N - 1)$.

- L'autocorrélation d'axes τ_8 est composée des vecteurs obtenus avec trois axes horizontaux :

$$\widehat{\mathcal{C}}_2[I](\tau_8) = \left(\widehat{\mathcal{C}}_2[I]\left(-\frac{N}{2}, \tau\right), \widehat{\mathcal{C}}_2[I](0, \tau), \widehat{\mathcal{C}}_2[I]\left(\frac{N}{2}, \tau\right) \right)^T \quad (4.36)$$

soit un vecteur de taille $3(2N - 1)$.

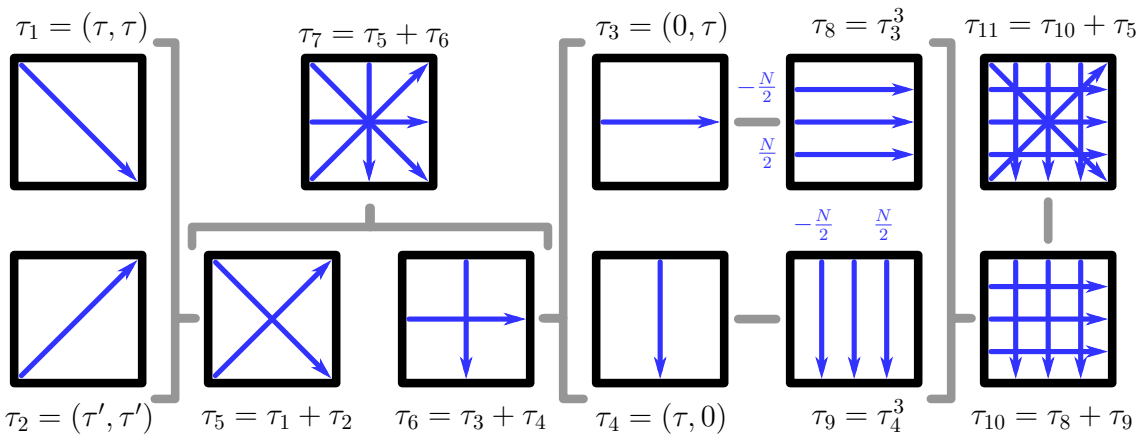


FIGURE 4.12 – Combinaisons d’axes issus des multicorrélations et sélectionnés pour définir les descripteurs de nos images : τ représente le retard (ou l’indice) qui varie et τ' signifie que la multicorrélation est calculée sur l’image pivotée de 90 degrés (pour obtenir la diagonale opposée). La face (*i.e.* le carré noir) représente les deux dernières dimensions de la multicorrélation. Par exemple, la bicorrélation (ordre trois) d’axe τ_3 d’une image I correspond au vecteur $\widehat{\mathcal{C}}_3[I](0, 0, 0, \tau)$. De manière similaire, le vecteur de la bicorrélation d’axe τ_6 sera composé des vecteurs de la bicorrélation calculée pour les axes τ_3 et τ_4 .

En totalité, onze combinaisons d’axes seront utilisées pour décrire les images de thrombi lors des simulations présentées au chapitre 5. Une autre alternative, plus gourmande en temps de calcul, serait de réduire les images de sortie de tailles $(2N - 1) \times (2N - 1)$ avec les méthodes de réduction de dimension présentées en section 2.3. Néanmoins, outre sa simplicité, l’avantage de choisir des axes fixes est que toutes les images seront décrites de manière similaire (contrairement aux méthodes de réduction de dimension qui pourraient projeter les images dans des espaces certes réduits mais de directions différentes).

4.2.4 Conclusion sur l’approche par statistiques

La section précédente exposait la construction des descripteurs des images de thrombi de notre base de données à partir d’une approche par ondelettes : le *scattering operator*. Dans cette section, nous nous sommes penchés sur une approche totalement différente, les statistiques d’ordre supérieur. Les ordres supérieurs à deux sont essentiellement utilisés en complément au second ordre, afin de résoudre des problèmes restés insolubles jusqu’alors, ou afin d’améliorer les solutions déjà apportées par les techniques classiques. Dans cette section, nous proposons d’étendre les définitions des multicorrélations des signaux monodimensionnels (*cf.* annexe D) aux images. L’objectif est donc d’extraire des attributs de nos images de thrombi, appelés descripteurs statistiques. La sous-section 4.2.2 a montré que le calcul des multicorrélations des images génère des descripteurs (tenseurs) de dimension bien plus grande que celle des images. Or, en section 2.3, nous avons vu que l’un des buts de l’extraction de descripteurs est de réduire la dimension des données. Pour cette raison, nous proposons en sous-section 4.2.3 de sélectionner seulement certains axes élémentaires des tenseurs des multicorrélations. A l’avenir, nous pourrions considérer des techniques de diagonalisation ou de contraction de tenseur pour réduire la dimension des multicorrélations. Dans la section suivante, nous suggérons d’associer les deux types de descripteurs envisagés, multirésolutions et statistiques pour caractériser les thrombi.

4.3 Combinaison du *scattering operator* et des statistiques d'ordre supérieur

En section 4.1, nous avons vu une première approche basée sur la théorie des ondelettes pour décrire, de manière plus succincte, les images de thrombi. Le *scattering operator* tente de caractériser une image en analysant son spectre (transformée de Fourier) pour déterminer si la texture de l'image comporte des éléments caractéristiques qui pourraient résumer au mieux l'image. L'idéal serait que ces éléments extraits soient propres aux images d'une même classe (catégorie, groupe) mais discriminant des autres classes (*cf. manifold untangling* en section 2.2.4). La section 4.2 propose une seconde approche qui vise à décrire les images à partir de leurs propriétés statistiques. Ces descripteurs statistiques, en s'intéressant à la distribution des pixels, devraient nous permettre d'obtenir une description des images différente de celle du *scattering operator*, qui est davantage orientée vers leur texture. Ces deux approches pourraient donc s'avérer complémentaires. Cette section succincte est axée sur la combinaison de ces approches et notamment sur la problématique de l'association de descripteurs de nature totalement différente. Cette section revient d'abord sur la structure des descripteurs des images obtenus par les deux approches, puis elle présente les formes de normalisation qui seront utilisées lors des simulations.

4.3.1 Structure des descripteurs

En résumé, nous disposons d'une base de données contenant des images échographiques et élastographiques collectées par l'Unité d'Echo-Doppler et de Médecine Vasculaire. Au chapitre 3, nous avons prétraité ces acquisitions (récupération des données patients et des paramètres d'acquisition, égalisation d'histogramme, conversion des élastographies) et extrait des imagerie carrées (de taille 32×32 ou 64×64). Ensuite, aux sections 4.1 et 4.2, nous proposons d'extraire des descripteurs de ces imagerie suivant deux approches distinctes mais potentiellement complémentaires : le *Scattering Operator* (SO) et les Statistiques d'Ordre Supérieur (SOS).

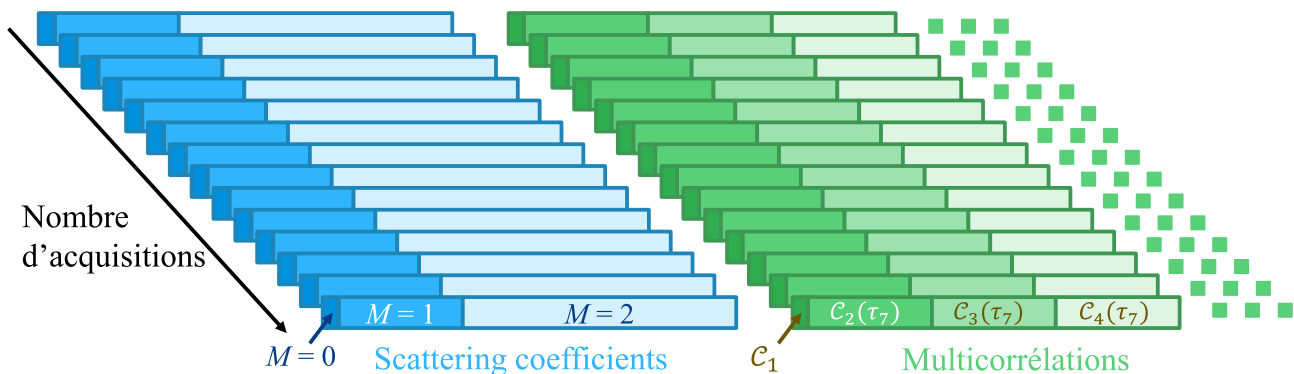


FIGURE 4.13 – Structure des descripteurs du SO et des Statistiques d'Ordre Supérieur (SOS) : pour chaque image de la base, nous disposons donc d'un vecteur issu du SO (nommé *scattering coefficients* d'ordre $M \in \{0, 1, 2\}$) et d'un autre issu des cumulants d'ordre $q \in \{2, 3, 4\}$. Pour ces deux vecteurs, on peut montrer que leur énergie décroît avec l'ordre.

Le SO peut être assimilé à un réseau de convolutions (voir figure 4.3) dont l'architecture dépend d'un certain nombre de paramètres : échelles, orientations, ordres de décomposition. Ce dernier paramètre correspond à la profondeur d'analyse en résolution des images et est donc lié au nombre de couches du réseau. L'ordre un, qui représente l'information basse fréquence, est constitué d'un seul coefficient : le coefficient d'approximation. La deuxième couche du réseau n'est autre que le résultat de transformées en ondelettes classiques suivi d'un module et d'un filtre passe-bas gaussien (et d'une étape d'agrégation si besoin). Le nombre de coefficients à l'ordre deux est le produit entre le nombre d'échelles et le nombre d'orientations. La particularité du SO par rapport à une transformée en ondelettes classique repose sur les ordres de décomposition supérieurs à deux : une nouvelle transformée en ondelettes (accompagné

du module et du passe-bas) est appliquée à chacun des modules des transformées en ondelettes de la couche du réseau précédente. Ce procédé, qui peut être répété à volonté, permet d'atteindre un niveau de détail de l'image examinée de plus en plus élevé. En pratique (et durant les simulations), nous sommes limités à l'ordre deux dont nous montrerons qu'il est bien souvent suffisant pour obtenir de bonnes performances. La figure 4.13 illustre la composition des descripteurs obtenus avec le SO. De plus, il est à noter que l'énergie des coefficients décroît de manière exponentielle avec l'ordre [22] (d'où l'intensité décroissante de la couleur bleue sur la figure 4.13).

En ce qui concerne les SOS, nous suggérons donc d'estimer les multicorrélations (cumulantes) des imagettes suivant plusieurs ordres et différentes combinaisons d'axes (*cf.* section 4.2.3). Les vecteurs descripteurs obtenus avec les SOS sont également schématisés sur la figure 4.13. De manière assez similaire au SO, les descripteurs statistiques ont une énergie qui décroît avec l'ordre des multicorrélations. Lors des simulations, nous tenterons notamment de déterminer quels axes et quels ordres sont les plus discriminants. Par leurs définitions, nous pouvons constater que ces deux types de descripteurs (SO et SOS) sont de nature différente et que, pour les associer, il faut donc certainement les transformer avant de procéder à l'étape de classification.

4.3.2 Uniformisation des descripteurs

En analyse des données, quelle que soit l'application, nous sommes souvent amenés à associer des informations de nature différente. Dans cette thèse, nous nous sommes penchés essentiellement sur les images des thrombi, mais nous aurions pu inclure certaines données patients¹ comme l'âge ou le taux de D-dimères dans le sang (*cf.* section 1.1.4). Pour comparer des données x de nature différente, nous proposons deux solutions parmi les plus usuelles :

0. Les descripteurs bruts (sans transformation) ;
1. La *standardisation* qui revient à la centrer (soustraction par sa moyenne \bar{x}) et à la réduire (division par son écart-type σ_x)

$$x \leftarrow \frac{x - \bar{x}}{\sigma_x} \quad (4.37)$$

2. La *normalisation min-max* qui consiste à ramener les données entre 0 et 1 :

$$x \leftarrow \frac{x - \min_x}{\max_x - \min_x} \quad (4.38)$$

L'utilisation des propositions ci-dessus soulève au moins deux autres questions : faut-il les appliquer sur l'ensemble des images (*i.e.* prendre la moyenne des descripteurs de toutes les images pour la standardisation) ou sur chaque image indépendamment ? Sachant que l'énergie décroît en augmentant l'ordre, ne serait-il pas judicieux de procéder à une uniformisation par ordre plutôt que par type de descripteurs ? Lors des simulations, nous comparerons les performances de chacune des deux propositions en les appliquant sur les descripteurs de :

- A. Chaque type sur l'ensemble des images ;
- B. Chaque ordre sur l'ensemble des images ;
- C. Chaque image et chaque type ;
- D. Chaque image et chaque ordre.

1. L'ajout de données biologiques comme le taux de D-dimères fait partie des pistes d'amélioration pour caractériser les thrombi et le choix de ces données est en cours d'étude au sein de l'Unité d'Echo-Doppler et de Médecine Vasculaire.

Outre le fait de pouvoir combiner les descripteurs multirésolutions et statistiques, l'uniformisation des descripteurs pourrait également, au même titre que l'égalisation d'histogramme, permettre de limiter l'influence des changements des paramètres d'acquisition entre les patients.

4.3.3 Conclusion sur la combinaison des deux approches

Dans cette courte section, nous proposons d'associer les deux approches des sections 4.1 et 4.2 (*scattering operator* et statistiques d'ordre supérieur). Ces deux approches fournissant des descripteurs de nature différente, nous envisageons donc huit manières (neuf avec le cas brut) d'uniformiser ces descripteurs : standardisation ou normalisation ; par type ou par ordre ; sur l'ensemble des images ou propre à chaque image.

4.4 Conclusion sur l'extraction de descripteurs multirésolutions et statistiques

Dans ce chapitre, nous avons détaillé les deux approches développées au cours de la thèse pour décrire les images extraites des acquisitions de thrombi veineux : le *scattering operator* et les statistiques d'ordre supérieur. La première citée, développée par Mallat [105] dans les années 2000, est une technique qui s'appuie sur une exploration multirésolution des signaux (*cf.* annexe C). Le *scattering operator* propose d'aller plus loin (en résolution) que les techniques de transformées en ondelettes classiques en s'inspirant des réseaux de convolution qui sont utilisés en apprentissage profond (*cf.* section 2.4.1.3). La principale différence est que, pour le *scattering operator*, les filtres de convolution sont fixés en amont (à partir des paramètres de l'algorithme) alors que ces filtres sont appris à partir d'une base d'apprentissage en ce qui concerne l'apprentissage profond. Par son réseau de convolution, le *scattering operator* permet de générer un espace de représentation des images qui est stable aux petites déformations et aux petites translations. Cette transformation de l'espace permet de conserver l'information discriminante des images (notamment sur leur texture) tout en réduisant leur dimension. Pour réduire encore davantage l'espace, nous avons vu qu'il est également possible de rendre cet espace invariant en échelle et en rotation. La seconde approche développée est une extension des outils statistiques classiques d'ordre deux (écart-type, variance, corrélation) aux ordres supérieurs. La littérature montre que l'utilisation de ces ordres supérieurs peut permettre une amélioration des performances, voire la résolution de problèmes insolubles à l'ordre deux. La section 4.2 introduit d'abord l'ensemble des notions mathématiques, et leurs propriétés, nécessaire pour comprendre leur utilisation sur des images. Parmi les quantités statistiques permettant de caractériser un signal, nous avons construit nos descripteurs statistiques à partir des cumulants temporels, appelés multicorrélations. Enfin, en section 4.3, nous proposons d'uniformiser les descripteurs multirésolutions (*scattering operator*) et statistiques (multicorrélations) puis de les combiner. En effet, ces deux types de descripteurs pourraient s'avérer complémentaires et ainsi améliorer la caractérisation des images de thrombi.

Chapitre 5

Simulations

Glossaire du chapitre 5

2D Deux Dimensions

ACP Analyse en Composantes Principales (ou *Principal Component Analysis* en anglais, PCA)

AME Apprentissage Multi-Echelle (ou multirésolution)

CLAHE Egalisation d'histogramme locale : *Contrast Limited Adaptive Histogram Equalization*

CS Classification Spectrale (*spectral clustering* en anglais)

DBSCAN Density-Based Spatial Clustering of Applications with Noise est un algorithme de partitionnement spatial à base de densité.

EP Embolie Pulmonaire

ID Identifiant

OPTICS Ordering Points to Identify the Clustering Structure est un algorithme de partitionnement spatial à base de densité.

SMA Classification par Sélection de Modèles Affines

SO *Scattering Operator* : réseau convolutif de transformée en ondelettes et d'opérateurs contractants

SOS Statistiques d'Ordre Supérieur (*High Order Statistics* en anglais, HOS)

t-SNE t-distributed Stochastic Neighbor Embedding

TVP Thrombose Veineuse Profonde

KTH-TIPS Base de textures créée par des membres de l'Institut royal de technologie de Stockholm (*Kungliga Tekniska Högskolan*) et contenant dix classes de matériaux observés sous différentes luminosités, orientations et échelles (*Textured under varying Illumination, Pose and Scale*) [56].

VP Vrais Positifs c'est-à-dire le nombre (ou le pourcentage) de patients malades classés correctement.

FN Faux Négatifs c'est-à-dire le nombre (ou le pourcentage) de patients malades classés comme normals.

VN Vrais Négatifs c'est-à-dire le nombre (ou le pourcentage) de patients sains classés comme normals.

FP Faux Positifs (ou fausse alarme) c'est-à-dire le nombre (ou le pourcentage) de patients sains classés comme malades.

Introduction

Si nous revenons à la chaîne de traitements d’images illustrée à la figure 2.4, le chapitre 1 a présenté les deux premiers blocs de la chaîne : le phénomène physique (la formation de thrombus veineux) et les acquisitions (échographies et élastographies avec le système Aplio 500). Le chapitre 3 a détaillé les deux blocs suivants à savoir la préparation des données et la segmentation (*i.e.* l’extraction des imagerie dans les régions d’intérêts). Le cinquième maillon de la chaîne, l’extraction de descripteurs, a fait l’objet du chapitre 4 où deux approches distinctes ont été proposées : le *Scattering Operator* (SO) et les Statistiques d’Ordre Supérieur (SOS). Dans cet ultime chapitre, nous proposons d’appliquer certains des algorithmes de classification présentés au chapitre 2 (SMA, k -moyennes, DBSCAN, CS) pour retrouver, à partir des descripteurs des images de thrombi, l’origine de leur formation et/ou la présence d’une Embolie Pulmonaire (EP). Néanmoins, avant d’appliquer les deux approches d’extraction de descripteurs et les algorithmes de classification sur la base de données de thrombi, nous proposons de régler certains de leurs paramètres et d’évaluer leurs performances sur une base de référence (*i.e.* avec des points de comparaison dans la littérature). L’ensemble des simulations menées sur cette base est résumé à la section 5.1. Ensuite, la section 5.2 propose de valider ces mêmes approches et algorithmes sur une base plus proche (en termes de type d’images) de celle des thrombi, la base des glandes salivaires. Enfin, en section 5.3, nous nous consacrerons aux images de thrombi afin de tenter de répondre à la problématique de notre projet (origine et/ou EP) à partir des résultats des simulations.

5.1 Validation et optimisation sur une base de référence

Dans cette section, nous proposons de tester les approches de caractérisation et de classification sur une base, dite de référence, dans le but de valider leur implémentation, d’optimiser certains de leurs paramètres clés et de comparer leurs performances à celles existantes dans la littérature. La sous-section 5.1.1 présente la base utilisée et l’intérêt de son utilisation vis-à-vis des bases de thrombi et de glandes salivaires. Les sous-sections 5.1.2 et 5.1.3 sont dédiées, respectivement, à l’application du SO et à celle des SOS pour classer les images de la base de référence. Ensuite, la sous-section 5.1.3 est consacrée à l’utilisation conjointe des deux approches d’extraction de descripteurs.

5.1.1 Intérêt et présentation de la base KTH-TIPS

La validation et l’optimisation des approches d’extraction de descripteurs (SO et SOS) sont difficilement réalisables directement sur les bases d’échographies et d’élastographies pour deux raisons principales : un manque de données et des classes de données pas forcément fiables à cent pourcent. L’optimisation paramétrique nécessite, pour être significative, que chaque classe soit clairement identifiée et comporte une cinquantaine (voire des centaines ou des milliers) de données. Or, que ce soit les thrombi ou les glandes salivaires, nous ne disposons pas encore de bases suffisamment riches en données et fiables en classes. Pour les glandes salivaires, il pourrait y avoir une classe intermédiaire composée des glandes normales suspectes et/ou des glandes atteintes du syndrome de Gougerot-Sjögren de stade 1 (la majorité des glandes étiquetées *Gougerot* sont de stade 3, le stade maximal). Quant aux thrombi, nous ne pouvons pas être certains que l’origine de leur formation et/ou leur évolution en EP influent sur leur structure ou que, si ces phénomènes influent, les modifications structurelles se traduisent sur leurs échographies et/ou leurs élastographies. Par conséquent, nous proposons de mener cette optimisation paramétrique et de valider l’implémentation des techniques d’extraction de descripteurs et de classification sur une base de référence : la base KTH-TIPS (Kungliga Tekniska Högskolan - Textured under varying Illumination, Pose and Scale).

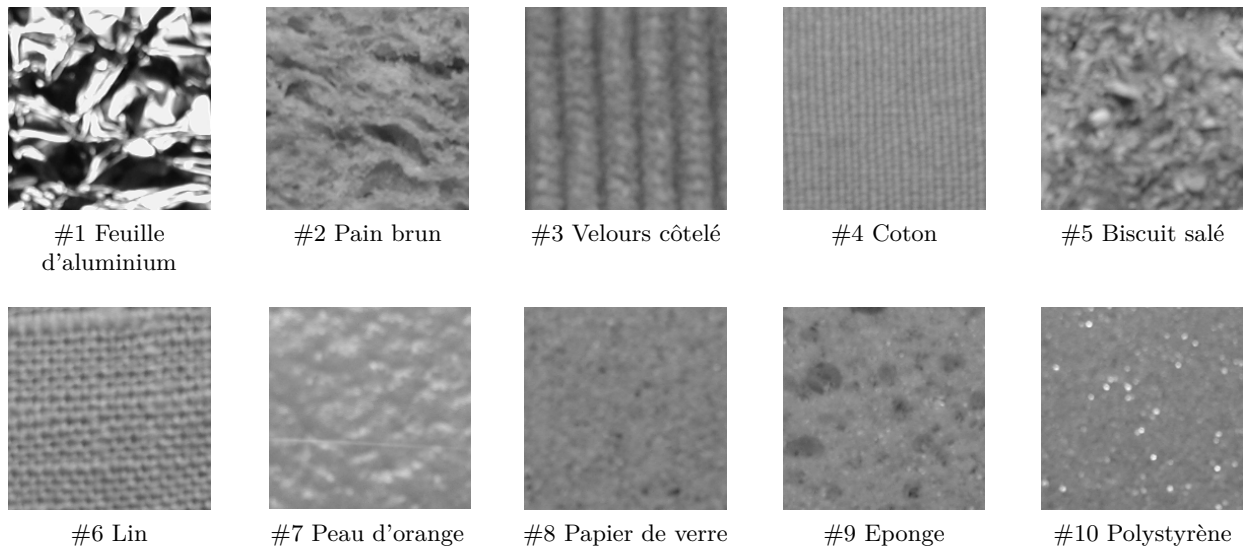


FIGURE 5.1 – Classes de textures présentes dans la base de données KTH-TIPS [56].

La base d'images texturées KTH-TIPS [56] a été créée par des membres de l'Institut royal de technologie de Stockholm (Kungliga Tekniska Högskolan ou KTH) à partir de différents matériaux observés sous différentes luminosités, orientations et échelles (Textured under varying Illumination, Pose and Scale ou TIPS). Cette base contient dix types de textures illustrées en Figure 5.1 et comportant 81 images chacun (810 images au total). Cette base est inspirée de la base CURET (Columbia-Utrecht Reflectance and Texture) [46] qui comprend soixante classes de textures. Etant donné que notre base de données de thrombus ne comporte pas plus de dix classes, le choix de la base KTH-TIPS semble plus judicieux que la base CURET. De plus, les résultats de nos simulations, sur cette base, pourront ensuite être comparés avec ceux existant dans la littérature [43, 99, 124, 151].

5.1.2 Scattering operator

Dans cette section, la base KTH-TIPS est utilisée pour régler les paramètres clés du SO (échelles, orientations et ordres de décomposition) et évaluer l'apport du logarithme, de l'Apprentissage Multi-Echelle (AME), des invariances en rotation et en échelle dans ses performances.

5.1.2.1 Sélection des paramètres clés

Le SO est un algorithme qui offre une certaine modulabilité qui lui permet de s'adapter au contexte (type de signal, taille du signal, échelle et orientation des éléments discriminants du signal). Par conséquent, certains paramètres du SO doivent être ajustés pour optimiser ses performances (*cf.* section 4.1.2). Pour ce faire, nous suggérons de mener une validation croisée des paramètres afin de sélectionner :

- la taille des imquettes N : 16×16 , 32×32 , 64×64 ou 128×128 ;
- le nombre d'échelles J : 2, 3, 4, 5 ou 6 ;
- le nombre d'orientations L : 1, 2, 4, 6 ou 8
- l'ordre maximal de décomposition M : 1 ou 2.

Parmi les techniques existantes, nous avons retenu la procédure de validation croisée nommée *v-fold cross-validation* en anglais [156]. Elle consiste à diviser chaque classe en v sous-ensembles puis de sélectionner un des v sous-ensembles de chaque classe pour former la base de test (appelée aussi base de validation). Les $v - 1$ autres ensembles constituent alors la base d'apprentissage. Les images de la base de test sont alors classées à partir des modèles générés avec la base d'apprentissage et le taux d'erreur de classement est alors calculé. L'opération est ensuite répétée $v - 1$ fois jusqu'à ce que chaque sous-ensemble ait été utilisé comme ensemble de test. Un schéma explicatif est visible en figure 5.2. La moyenne des v erreurs est enfin calculée pour estimer l'erreur de classification. Ces v opérations doivent être réitérées pour chaque jeu de paramètres afin de déterminer le jeu qui minimise l'erreur de classification. En pratique, comme chacune des dix classes contient 81 images, la base est partitionnée en neuf sous-ensembles ($v = 9$).

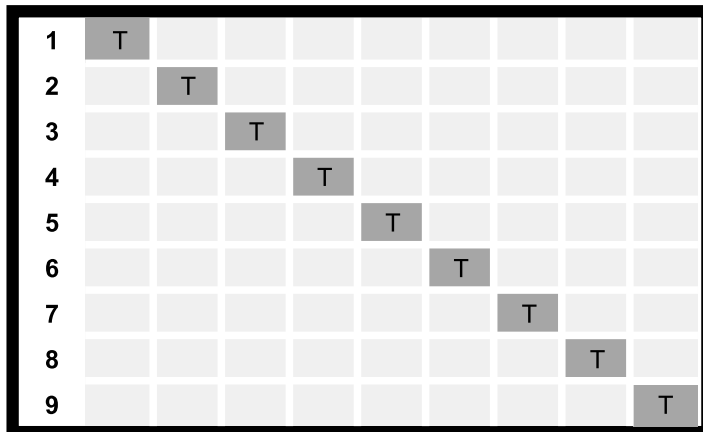


FIGURE 5.2 – Procédure de validation croisée en neuf sous-ensembles : à chaque tour, un des neuf sous-ensembles forme la base de validation et les huit sous-ensembles restants constituent la base d'apprentissage.

La procédure de validation croisée décrite précédemment nécessite donc de modéliser chaque classe avec les $v - 1$ sous-ensembles de la base d'apprentissage. Pour ce faire, nous utilisons un classifieur linéaire simple et complémentaire du SO : la classification par Sélection de Modèles Affines (SMA) (cf. section 2.4.1.1) qui utilise l'Analyse en Composantes Principales (ACP) pour réduire la dimension des données. L'objectif de la SMA est d'obtenir le résumé le plus pertinent possible des données issues des *scattering* coefficients pour chacune des dix classes de la base d'apprentissage. L'ACP analyse les corrélations entre les données pour obtenir un petit nombre de nouvelles variables non corrélées à forte variance (i.e. à fort pouvoir informatif). Le choix du nombre d de ces nouvelles variables est à définir. Une fois que chaque classe est modélisée, les images de la base de test sont étiquetées à la classe du modèle le plus similaire. D'un point de vue pratique, le paramètre d (dimension de l'espace affine d'une classe), qui doit être optimisé, prend les valeurs suivantes durant nos simulations : 4, 8, 12, 14, 16, 18, 20, 24, 30, 36, 42, 50, 64, 96 ou 128. En fonction des paramètres du SO, les valeurs de d supérieures au nombre de descripteurs ne sont pas prises en compte (la dimension de l'espace projeté ne peut pas être supérieure à celle de l'espace des descripteurs).

La figure 5.3 présente les résultats de la validation croisée pour les différents paramètres. Chaque case du tableau représente le taux de mauvaise classification pour un nombre d'orientations, d'échelles, d'ordres, une dimension des espaces affines de projection et une taille d'image donnés. Le taux d'erreur est considéré comme satisfaisant s'il est inférieur à 10 % (i.e. si la case n'est pas de couleur bleue). Globalement, on peut voir que plus les nombres d'ordres, d'échelles, d'orientations et de pixels sont élevés, plus le SO est performant pour discriminer les dix classes de textures. On voit notamment que les images doivent être de taille supérieure ou égale à 64×64 pour obtenir des résultats satisfaisants. Il faut ainsi deux ordres de décomposition et au moins trois échelles et quatre orientations. Enfin la dimension de l'espace affine peut varier entre 12 et 128 suivant le nombre de descripteurs sans dégrader les résultats.

La figure 5.4 résume les paramètres du SO les plus performants pour 64×64 et 128×128 pixels. Pour 64×64 , les meilleurs résultats sont obtenus pour quatre ou cinq échelles et six ou huit orientations et ils oscillent entre 2 et 4 % d'erreur moyenne. Pour 128×128 , ils sont inférieurs à 1.5 % si le SO est appliqué avec au moins cinq échelles et quatre orientations. Hormis précision indiquant le contraire, les prochaines simulations sont effectuées avec huit orientations ($L = 8$), cinq échelles ($J = 5$) et deux ordres de décomposition ($M = 2$) car ils semblent légèrement meilleurs quelle que soit la taille de l'image.

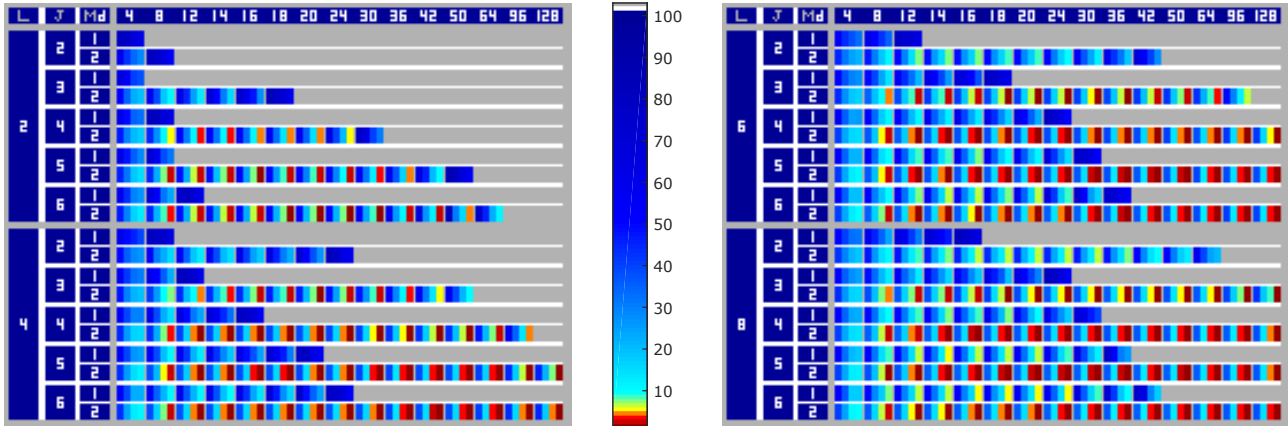


FIGURE 5.3 – Procédure de validation croisée appliquée à la base KTH-TIPS [46] : chaque case du tableau correspond au pourcentage d'erreur moyen obtenu pour chaque orientation L , échelle J , ordre M , dimension d des espaces affines et taille d'images N (une case est divisée en quatre barres correspondant respectivement aux quatre tailles d'image possibles 16×16 , 32×32 , 64×64 ou 128×128). Les cases grisées signifient que le nombre de descripteurs est inférieur à la dimension d .

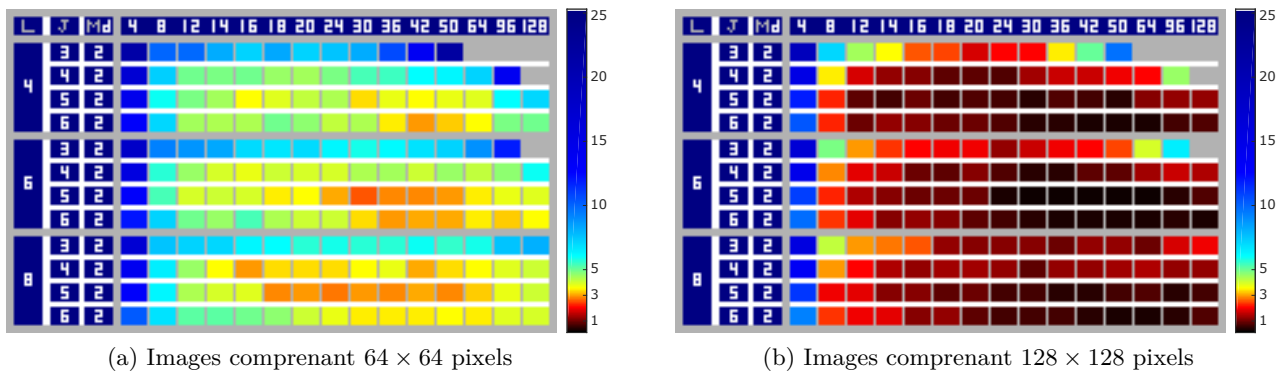


FIGURE 5.4 – Pourcentage d'erreur moyen pour les paramètres les plus performants : chaque case du tableau correspond au pourcentage d'erreur moyen obtenu pour chaque orientation L , échelle J , ordre M , dimension d des espaces affines.

5.1.2.2 Logarithme, invariances et apprentissage multi-échelle

La figure 5.5 donne les taux moyens de bonne classification et leurs écart-types en fonction des propriétés du SO et de la taille de la base d'apprentissage. Lorsque la base d'apprentissage est suffisamment grande (20 ou 40 images sur les 81 images que composent chaque classe), la SMA appliquée sur les *scattering* coefficients donne de très bons résultats quelle que soit la méthode utilisée (plus de 95 % de bonne classification). Avec un huitième des images consacré à l'apprentissage (10 image par classes), le taux de classification diminue aux alentours des 90 %, ce qui reste encore très bon. En divisant encore par deux la taille de la base d'apprentissage, le pourcentage de classification reste satisfaisant mais chute à environ 70 % si l'on utilise l'algorithme classique du SO (stabilité aux petites déformations et invariance en translation). Toutefois, on peut voir qu'en rajoutant certaines propriétés au SO (logarithme, invariances en rotation et en échelle), il est possible d'atteindre les 80 % de bonne classification.

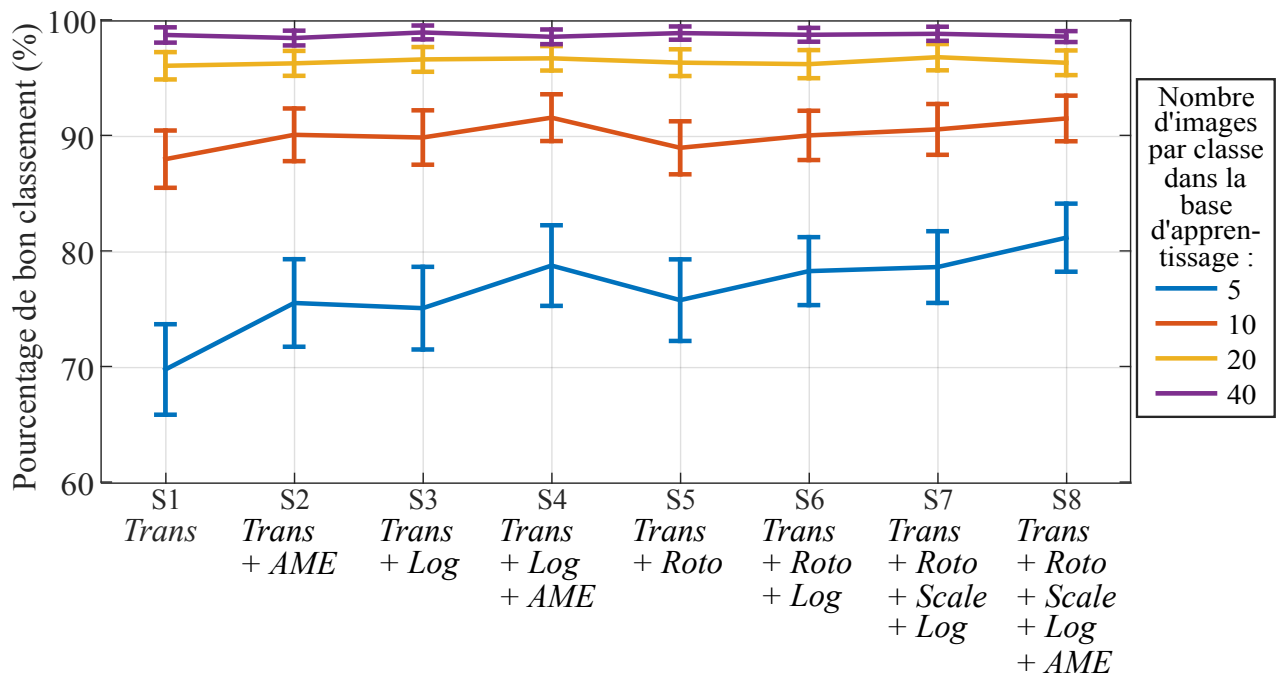


FIGURE 5.5 – Moyenne et écart-type (sur 100 partitions différentes et pour quatre proportions d'apprentissage/test) de la base KTH-TIPS [46] en fonction des propriétés du SO : logarithme des coefficients (*Log*), AME, invariances en translation (*Trans*), en échelle (*Scale*) et en rotation (*Roto*).

L'Apprentissage Multi-Echelle (AME) permet d'agrandir la base d'apprentissage en créant des répliques de chaque image d'apprentissage à des échelles différentes. Lorsque la base d'apprentissage contient peu d'images, (5 ou 10 images par classes), le fait de quadrupler sa taille rend le SO (suivi de la SMA) plus performant. De même, la classification est plus efficace lorsque l'on applique le logarithme avant le moyennage spatial (agrégation de la sortie du filtre passe-bas appliqué à chaque nœud du réseau de convolutions du SO). Le logarithme permet de linéariser la décroissance exponentielle des *scattering* coefficients en fonction de l'ordre de décomposition. L'ajout des invariances en rotation et en échelle améliorent également les performances du SO sur la base KTH-TIPS mais de manière moins significative que le logarithme et/ou l'AME.

C1	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10
#1	69	10	0	0	4	0	1	3	11	2
#2	0	71	0	0	3	0	3	8	13	2
#3	0	2	79	9	1	1	1	3	3	1
#4	0	1	4	65	0	3	0	17	4	7
#5	0	10	0	0	74	0	2	2	12	0
#6	0	2	2	4	0	68	3	9	2	10
#7	0	7	0	0	0	1	72	6	7	6
#8	0	2	0	2	0	0	2	64	9	21
#9	0	10	0	0	2	0	3	19	62	4
#10	0	0	0	0	0	0	2	19	3	75

(a) Invariance en translation et logarithme (S1)

C8	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10
#1	95	2	0	0	2	0	0	0	1	0
#2	1	82	0	0	6	0	1	2	8	0
#3	0	1	86	6	0	5	1	0	0	0
#4	0	0	3	78	0	10	1	7	0	1
#5	1	12	0	0	81	0	0	0	5	0
#6	0	0	5	22	0	71	0	1	0	0
#7	0	3	0	1	0	0	89	3	3	0
#8	0	0	0	3	0	0	1	74	10	11
#9	0	5	0	0	2	0	1	11	78	3
#10	0	1	0	0	0	0	1	12	6	80

(b) Apprentissage multi-échelles, logarithme, invariance en translation, en échelle et en rotation (S8)

Tableau 5.1 – Matrices de confusion de la classification par SMA avec cinq images par classes dans la base d'apprentissage pour illustrer l'intérêt de l'AME.

Les tableaux 5.1.a et 5.1.b comparent deux matrices de confusion moyenne obtenues avec deux types d'utilisation du SO. Chaque ligne de la matrice de confusion correspond à une classe de vérité alors que chaque colonne représente une classe prédite. En sommant chaque valeur d'une ligne, on obtient bien 100 % des images de la classe de vérité correspondante. Globalement, ces deux matrices montrent que l'ajout des différentes techniques énumérées précédemment (*AME + Log + Roto + Scale*) permet de diminuer grandement la confusion interclasse et donc d'améliorer la classification. On peut voir que la classe *Feuille d'aluminium* (#1) est bien classée 95 %. *A contrario*, la classe *Lin* (#6) est la moins bien classée (71 %) car 22 % de ces images sont confondues avec la classe *Coton* (#4).

Les résultats de nos simulations sont cohérents avec ceux obtenus par les auteurs de [151] affichés dans le tableau 5.2. Pour une petite base d'apprentissage, l'ajout des invariances en rotation, du logarithme, de l'invariance en échelle puis de l'apprentissage multi-échelles permettent d'améliorer graduellement le taux de classification. De plus, on peut noter que le scattering operator est au moins aussi performant (avec 40 images par classe dans la base d'apprentissage) que les trois autres techniques décrites et appliquées dans [151] :

- Log Gaussian Cox (COX) [124]
- Basic Image Features (BIF) [43]
- Sorted Random Projection (SRP) [99]

Train size	5	20	40
COX [124]	80.2 ± 2.2	92.4 ± 1.1	95.7 ± 0.5
BIF [43]	-	-	98.5
SRP [99]	-	-	99.3
Trans	69.1 ± 3.5	94.8 ± 1.3	98.0 ± 0.8
+ Roto	69.5 ± 3.6	94.9 ± 1.4	98.3 ± 0.9
+ Log	76.2 ± 3.3	96.0 ± 1.1	98.8 ± 0.7
+ Scale	77.8 ± 3.6	97.4 ± 1.0	99.2 ± 0.6
+ AME	84.3 ± 3.1	98.3 ± 0.9	99.4 ± 0.4

 Tableau 5.2 – Taux (moyennes et écart-types) de bonne classification sur la base KTH-TIPS pour trois proportions d'apprentissage/test (nombre d'images dans la base d'apprentissage par classe) publiés dans [151] : les cinq dernières lignes correspondent aux résultats obtenus avec le SO classique (*Trans*) puis en ajoutant au fur et à mesure l'invariance en rotation (*Roto*), le logarithme (*Log*), l'invariance en échelle (*Scale*) et l'AME.

Ces trois techniques visent, comme le SO à construire une nouvelle représentation qui est invariante en translation et en rotation. La première méthode COX caractérise la texture par des points clés qui sont déterminés à partir d'un modèle statistique gaussien dans différentes régions circulaires des images. La technique BIF représente les textures par des descripteurs estimés à partir des réponses à un banc de filtres dérivateurs gaussiens. Le troisième algorithme SRP décrit également la texture en appliquant des bancs de filtres orientés puis utilise un algorithme pour trier et projeter les réponses des différents filtres.

5.1.2.3 Classification non supervisée et paramètres du SO

Les deux sous-sections précédentes ont été consacrées à la classification supervisée des descripteurs multirésolutions du SO dans le but de restreindre au maximum le paramétrage de l'algorithme. En sous-section 5.1.2.1, nous avons vu que deux ordres de décomposition suffisent, qu'il faut quatre à huit orientations et quatre à six échelles. La sous-section 5.1.2.2 a montré que l'ajout du logarithme dans l'architecture du SO (et l'apprentissage multi-échelle) permettait d'améliorer les performances du SO mais que l'apport des invariances en rotation et en échelle n'était pas aussi flagrant sur cette base de test. Cette sous-section, consacrée à la classification non supervisée, propose de s'interroger sur les trois points suivants :

- Quelles sont les performances de l'algorithme des k -moyennes et de la classification spectrale sur les descripteurs du SO ?
- L'ajout du logarithme et des invariances (rotation, échelle) permet-il d'améliorer les performances de classification ?
- La construction des vecteurs descripteurs de chaque image à partir de plusieurs jeux de paramètres du SO (au lieu d'un seul jeu) permet-il d'accroître également les performances de classification ?

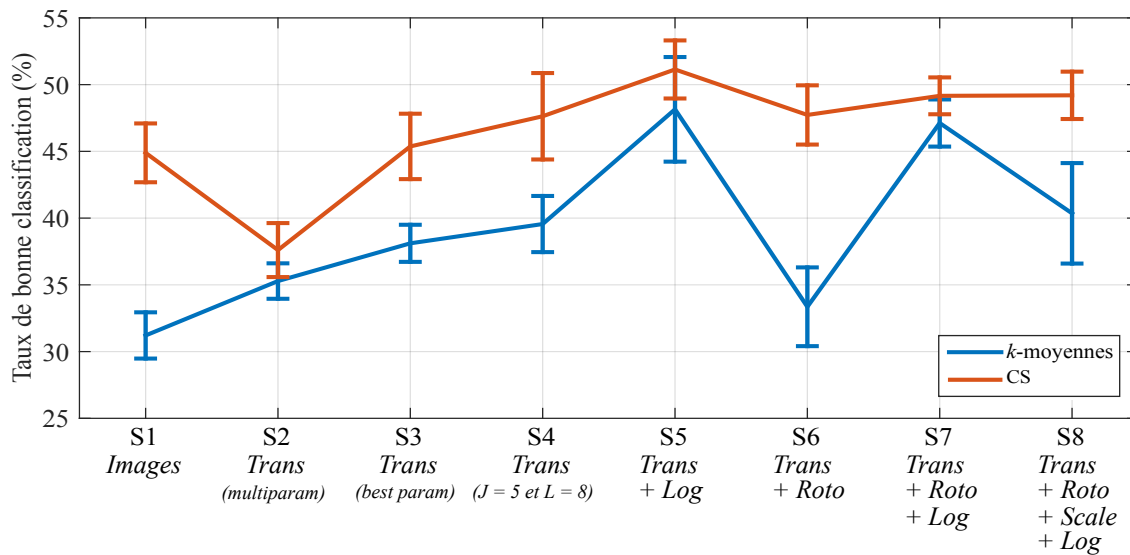


FIGURE 5.6 – Evaluation du SO avec deux algorithmes de classification non supervisée : les k -moyennes et la Classification Spectrale (CS). La simulation $S1$ consiste à appliquer les deux algorithmes de classification directement sur les images elles-mêmes. Dans les simulations $S2$, $S3$ et $S4$, ils sont appliqués aux *scattering* coefficients calculés avec les paramétrages respectifs suivants : tous les paramètres J et L de la procédure de validation croisée décrite en sous-section 5.1.2.1 (*multiparam*), les paramètres les plus optimaux $J = [4, 5, 6]$ et $[L = 4, 6, 8]$ (*best param*) et le couple de paramètres sélectionné pour les autres simulations ($J = 5$ et $L = 8$). Ensuite, les simulations $S5$, $S6$, $S7$ et $S8$ sont réalisées en ajoutant les propriétés suivantes : le logarithme (*Log*) et les invariances en rotation (*Roto*) et en échelle (*Scale*).

La figure 5.6 donne la moyenne et l'écart type des taux de bonne classification sur 10 tests (*i.e.* 10 initialisations aléatoires). La simulation $S1$ compare les performances de l'algorithme des k -moyennes et de la CS (avec les k -moyennes) appliqués directement sur les images. Pour les autres simulations, ces deux algorithmes de classification sont appliqués sur les descripteurs multirésolutions des images calculés avec différents paramétrage du SO. Que ce soit sur les images ou sur les descripteurs, on peut constater que la CS surpasse, parfois largement, les k -moyennes.

En ce qui concerne l'optimisation du SO, les simulations $S2$, $S3$ et $S4$ indiquent qu'il est plus judicieux d'utiliser un seul jeu de paramètres (ici $J = 5$ et $L = 8$) plutôt que de combiner les coefficients obtenus avec plusieurs jeux de paramètres. Ensuite, les simulations $S5$, $S6$, $S7$ et $S8$ montrent que l'utilisation du logarithme améliore les performances du SO (notamment lorsqu'il est associé aux k -moyennes mais que l'ajout des invariances en rotation et en échelle n'est pas forcément nécessaire (taux équivalents voire inférieurs).

Enfin, les matrices de confusion moyennes de la simulation $S5$ (translation et logarithme), affichées aux tableaux 5.3.a (k -moyennes) et 5.3.b (CS), signalent que les classes *Papier de verre* (#8) et *Eponge* (#9) font chuter le taux de classification. Les autres types de textures sont mieux classés (quasiment 50 %) voire même très bien classés (plus de 70 %) pour les *Feuille d'aluminium* (#1) et les *Peaux d'orange* (#7). Les résultats sont donc satisfaisants mais restent néanmoins inférieurs à ceux de la classification supervisée (par SMA).

S5	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10
#1	95	0	0	0	0	0	0	5	0	0
#2	0	53	0	0	34	0	0	0	13	0
#3	0	4	42	11	3	17	12	0	4	6
#4	0	0	13	38	0	15	3	12	3	16
#5	12	11	0	0	58	0	0	0	19	0
#6	0	2	0	0	0	50	5	5	1	36
#7	0	12	10	1	0	0	59	5	6	5
#8	0	3	0	17	0	15	13	17	2	33
#9	0	28	1	5	6	5	36	6	11	2
#10	0	2	0	1	0	19	5	15	0	58

 (a) k -moyennes : $48.15 \pm 3.92 \%$

S5	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10
#1	87	1	0	0	2	0	0	11	0	0
#2	0	49	0	0	27	0	0	1	23	0
#3	0	2	62	3	2	18	1	0	6	6
#4	0	0	16	47	0	16	2	7	1	11
#5	0	21	0	0	50	0	0	0	29	0
#6	0	2	2	2	0	52	9	4	3	26
#7	0	5	12	2	0	0	71	7	1	1
#8	0	2	0	18	0	16	15	18	2	29
#9	0	23	2	4	2	6	38	8	15	1
#10	0	2	0	3	0	19	5	9	4	59

 (b) Classification spectrale : $51.14 \pm 2.18 \%$

 Tableau 5.3 – Matrices de confusion de la classification des *scattering* coefficients (invariance en translation et logarithme) par k -moyennes et par CS (correspond à la simulation $S5$ en figure 5.6).

5.1.2.4 Bilan sur les descripteurs multirésolutions

Les simulations effectuées sur la base KTH-TIPS composée d'images texturées a permis de mettre en lumière certaines tendances sur l'application du *scattering operator*. En ce qui concerne son paramétrage, la procédure de validation croisée a permis de sélectionner les meilleurs jeux de paramètres (deux ordres de décomposition, au moins quatre orientations, au moins quatre échelles) et a montré qu'il était préférable que les imagerie soient les plus grandes possibles. Les différentes simulations ont aussi indiqué que l'ajout des invariances en échelle et/ou en rotation donne des résultats globalement équivalents (supérieurs en supervisé, inférieurs en non supervisé) à ceux acquis avec l'invariance en translation seule. Nous avons également observé que le logarithme, en linéarisant l'énergie en fonction des ordres de décomposition, renforce les informations hautes fréquences (*i.e.* de textures) contenues dans les ordres supérieurs (ordre deux notamment) et accroît considérablement les performances du *scattering operator*. De plus, l'apprentissage multi-échelle pourrait s'avérer pertinent pour classer de manière supervisée la base de données de thrombi veineux.

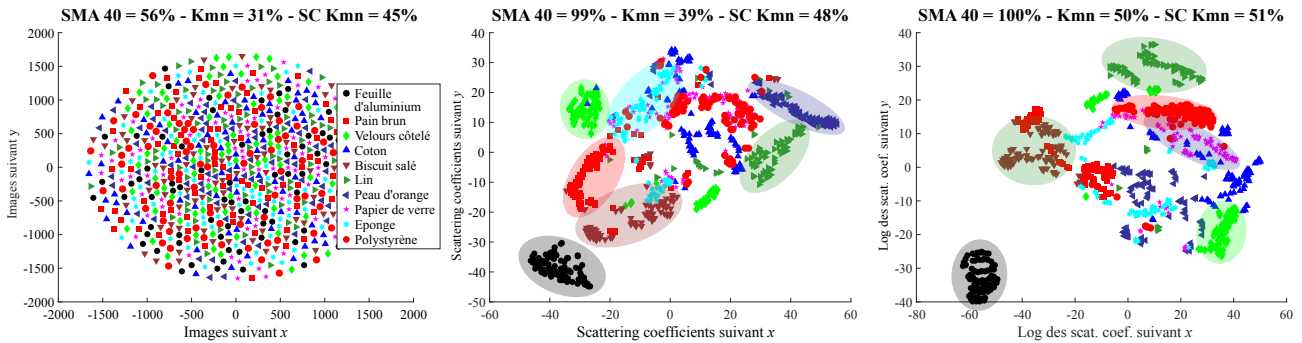


FIGURE 5.7 – *Manifold untangling* du *scattering operator* par la méthode t-distributed Stochastic Neighbor Embedding (t-SNE) présentée en section 2.3.3 : les descripteurs multirésolutions sont projetés dans un espace à deux dimensions qui préserve les distances (chaque point représente une image de la base KTH-TIPS) : le titre de chaque espace 2D indique les taux de bonne classification par SMA (50 % dédié à l'apprentissage), par k -moyennes (Kmn) et par classification spectrale ($SC Kmn$).

Sur cette base de données, la classification supervisée (au moins cinq images par classe consacrées à l'apprentissage) donne de bien meilleurs résultats que les approches non supervisées considérées dans cette section (k -moyennes ou classification spectrale). Le partitionnement par k -moyennes s'avère plus performant dans le domaine spectral (*i.e.* classification spectrale) que dans le domaine spatial (directement sur les images) ou dans celui des ondelettes (sur les *scattering* coefficients). La figure 5.7 illustre le *manifold untangling*, c'est-à-dire la déformation de la surface des données, (*cf.* section 2.2.4), avec la méthode t-SNE qui permet de réduire la dimension de l'espace de représentation en préservant les distances (*cf.* section 2.3.3). Cet outil montre que, comparé à l'espace de départ (les images brutes), l'espace des descripteurs multirésolutions fournit une nouvelle représentation des images dans laquelle les classes d'images sont plus simplement séparables (#1 *Feuille d'aluminium*, #2 *Pain*, #3 *Velours*, #5 *Biscuit*, #6 *Lin*, #7 *Peau d'orange*, #9 *Eponge*) bien que certaines classes soient encore entremêlées (#4 *Coton*, #8 *Papier de verre*, #10 *Polystyrène*). Cet indicateur 2D est plutôt cohérent avec les résultats de classification non supervisée (environ 50 %) dans le sens où, si la classe *aluminium* (#1) est bien isolée, les frontières entre les autres classes semblent moins nettes, ce qui rend la classification par partitionnement plus compliquée.

5.1.3 Multicorrélations

Dans cette section, les images de la base KTH-TIPS sont décrites par les multicorrélations d'ordre supérieur. Dans un premier temps, nous nous interrogerons sur l'ordre maximal à atteindre pour décrire au mieux les images puis, dans un second temps, nous étudierons le pouvoir discriminant des descripteurs statistiques en fonction des axes des cumulants estimés (*cf.* section 4.2.3).

5.1.3.1 Ordre maximal

Chaque image de la base KTH-TIPS est représentée par un vecteur comprenant l'ensemble des valeurs des multicorrélations. Celles-ci sont calculées de l'ordre un à l'ordre six et suivant l'axe diagonal τ_1 . La figure 5.8 donne les taux de classification supervisée (SMA) et non supervisée (k -moyennes, DBSCAN et CS), en fonction de l'ordre des multicorrélations. Les résultats de la première colonne comparent les performances de classification des descripteurs statistiques par ordre alors que les deux autres colonnes évaluent l'apport des ordres supérieurs aux multicorrélations d'ordre deux (deuxième colonne) ou d'ordres un et deux (troisième colonne).

En supervisé, trois proportions de la base d'apprentissage ont été considérées (50 %, 25 % et 6.25 %). Avec au moins 25 % des images consacrées à l'apprentissage, les taux de bonne classification oscillent entre 60 % et 90 %, ce qui est plutôt satisfaisant étant donné qu'il y a dix classes. Nous pouvons observer que l'ordre trois, pris isolément, est l'ordre le plus discriminant pour classer les images de cette base. Le fait de combiner les ordres des multicorrélations permet d'accroître les performances : le vecteur descripteur formé des ordres deux, trois et quatre améliore d'au moins 10 % les taux de classification de ces trois ordres pris séparément. Les ordres cinq et six n'apportent pas de réel gain de performance et ont même tendance à dégrader légèrement les résultats : ces deux ordres ajoutent probablement de l'information redondante (ou peu discriminante) qui viennent « noyer » l'information pertinente des ordres inférieurs. Avec 6.25 % d'images dédiées à l'apprentissage, nous obtenons des taux de classification proches de ceux de la classification non supervisée (entre 40 et 60 %), ce qui reste bien mieux qu'une classification aléatoire des images (10 %) mais qui n'est pas assez satisfaisant.

La classification non supervisée a été menée de quatre façons différentes :

1. Par un algorithme de partitionnement (deuxième ligne de la figure 5.8) :
 - (a) k -moyennes ;
 - (b) DBSCAN ;
2. Par classification spectrale où le partitionnement est réalisé par (troisième ligne) :
 - (a) k -moyennes ;
 - (b) DBSCAN.

L'algorithme des k -moyennes (ou *k-means*, voir sous-section 2.4.2.1) est appliqué avec dix classes et dix initialisations aléatoires des dix *clusters*. Les taux de bonne classification par cet algorithme représente donc la moyenne des taux obtenus sur les dix tests.

Durant nos simulations, l'algorithme Density-Based Spatial Clustering of Applications with Noise (DBSCAN) est utilisé avec les paramètres suivants (*cf.* sous-section 2.4.2.2) :

- Le nombre de points minimum v nécessaires pour définir un *cluster* prend l'une des dix valeurs suivantes : 3, 5, 7, 8, 9, 10, 11, 12, 15 et 20¹. Le taux de classification affiché en figure 5.8 pour le DBSCAN est le taux maximal obtenu avec ce jeu de valeurs de v .
- Le rayon de voisinage r est défini, d'une manière proche de celle du paramètre de dispersion en classification spectrale (*cf.* annexe E), comme la médiane des distances entre chaque descripteur et son v -ième voisin le plus proche. Cette astuce permet notamment d'automatiser le choix du rayon r en fonction du paramètre v ².

Pour s'adapter à la base KTH-TIPS qui contient dix classes, deux étapes ont été ajoutées à la fin de l'algorithme DBSCAN pour que ces performances soient davantage comparables à celles des k -moyennes :

1. Les points qui ne sont pas classés dans les dix plus gros *clusters* sont considérés comme aberrants ;
2. Les points aberrants sont rattachés au même *cluster* de leur voisin (non aberrant) le plus proche.

1. Le choix de ces dix valeurs a été défini de manière empirique sachant que les expérimentations préliminaires ont montré qu'un nombre v trop important (supérieur à 20 pour la base KTH-TIPS) classait systématiquement l'ensemble des images comme *aberrantes*.

2. Comme pour le choix de v , des tests en amont ont été réalisés pour définir le rayon r le plus pertinent possible. Le fait de définir le rayon r en fonction de v , et donc pas de manière indépendante, permet de simplifier l'évaluation de l'algorithme (un seul paramètre à faire varier) mais aussi d'adapter le rayon r aux données elles-mêmes puisqu'il sera proportionnel à la distance qui sépare les données. Le choix de la médiane des distances s'est avéré expérimentalement plus judicieux que le quartile ou la moyenne.

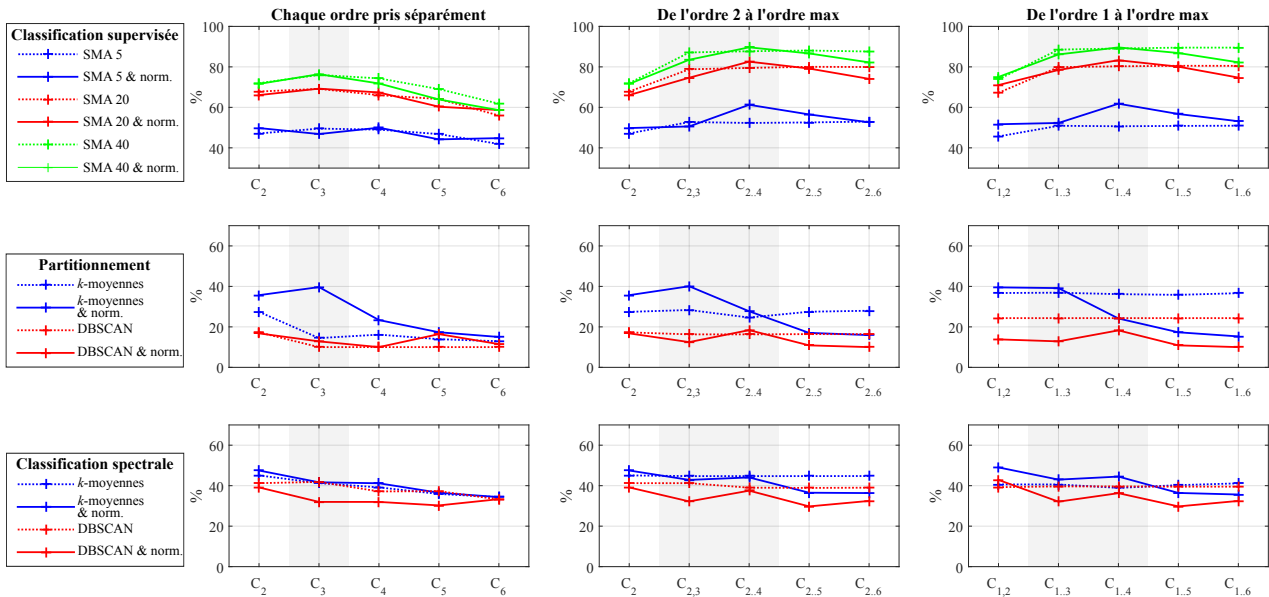


FIGURE 5.8 – Pouvoir discriminant des multicorrélations en fonction de leur ordre : taux de bonne classification en fonction des ordres des multicorrélations. Les traits pleins (par rapport aux traits pointillés) indiquent que les cumulants d'ordre $q > 2$ des images I ont été normalisés par la puissance q de leur écart-type σ_I (et par l'écart-type à l'ordre 2). Les lignes correspondent aux résultats respectifs de la classification supervisée par SMA (taux moyen sur dix tests) pour trois proportions de la base d'apprentissage (5, 20 et 40 images par classes soit 6.25 %, 25 % et 50 % de la base totale), des algorithmes de partitionnement k -moyennes (taux moyen sur dix test) et DBSCAN (taux maximal pour dix valeurs du paramètre *Nombre de points minimum nécessaire pour créer un cluster*). La première colonne compare les taux de classification pour chaque ordre pris séparément. En deuxième et troisième colonnes, les cumulants d'ordres différents sont associés de manière croissante (ordre deux, ordres deux et trois, ordres deux, trois et quatre, *etc.*), la différence étant que l'ordre un (moyenne) a été ajouté aux vecteurs descripteurs pour la troisième colonne.

Globalement, la figure 5.8 semble indiquer que la classification spectrale donne de meilleurs résultats que l'application directe des algorithmes de partitionnement (k -moyennes ou DBSCAN). L'algorithme des k -moyennes paraît également plus performant que le DBSCAN pour classer les descripteurs statistiques. Comme en supervisé, les ordres deux et trois sont les ordres les plus discriminants pour classer les images de la base KTH-TIPS. Le saut de performance dû à la fusion des descripteurs pour plusieurs ordres est moins flagrant en non supervisé où les performances chûtent au-delà de l'ordre trois (en particulier en l'absence de classification spectrale). Nous pouvons noter que l'ordre un (moyenne), lorsqu'il est ajouté aux descripteurs non normalisés (pointillé), joue un rôle prépondérant dans le classement des images par les k -moyennes (troisième colonne et deuxième ligne de la figure 5.8). Hormis ce cas particulier, la normalisation des descripteurs par l'écart-type à la puissance adéquate permet généralement d'améliorer les résultats de classification (les traits pleins sont, le plus souvent, au-dessus des traits pointillés).

A présent, nous allons nous intéresser au *manifold untangling* comme nous l'avons fait pour le *scattering operator*. La figure 5.9 représente les descripteurs statistiques projetés en deux dimensions par la méthode t-SNE. Comparée à la projection des images en deux dimensions (2D), la projection des descripteurs statistiques indique que les images sont plus facilement séparables dans l'espace des multicorrélations que dans le domaine d'origine (domaine spatial des images). En ajoutant les ordres trois et quatre à l'ordre deux, nous observons que certaines classes semblent se détacher davantage du nuage de points. Ces constatations se traduisent aussi en chiffre, et tout spécialement en supervisé où le taux de classification passe de 56 % sur les images à 90 % sur les multicorrélations d'ordres deux, trois et quatre. En non supervisé, le gain apporté par le changement d'espace (des images vers les statistiques) est pour l'instant moins flagrant. Dans la prochaine sous-section, nous allons considérer

d'autres axes des multicorrélations par rapport aux descripteurs utilisés dans cette sous-section (obtenu avec l'axe diagonal τ_1) et observer si cela permet d'améliorer les performances de classification.

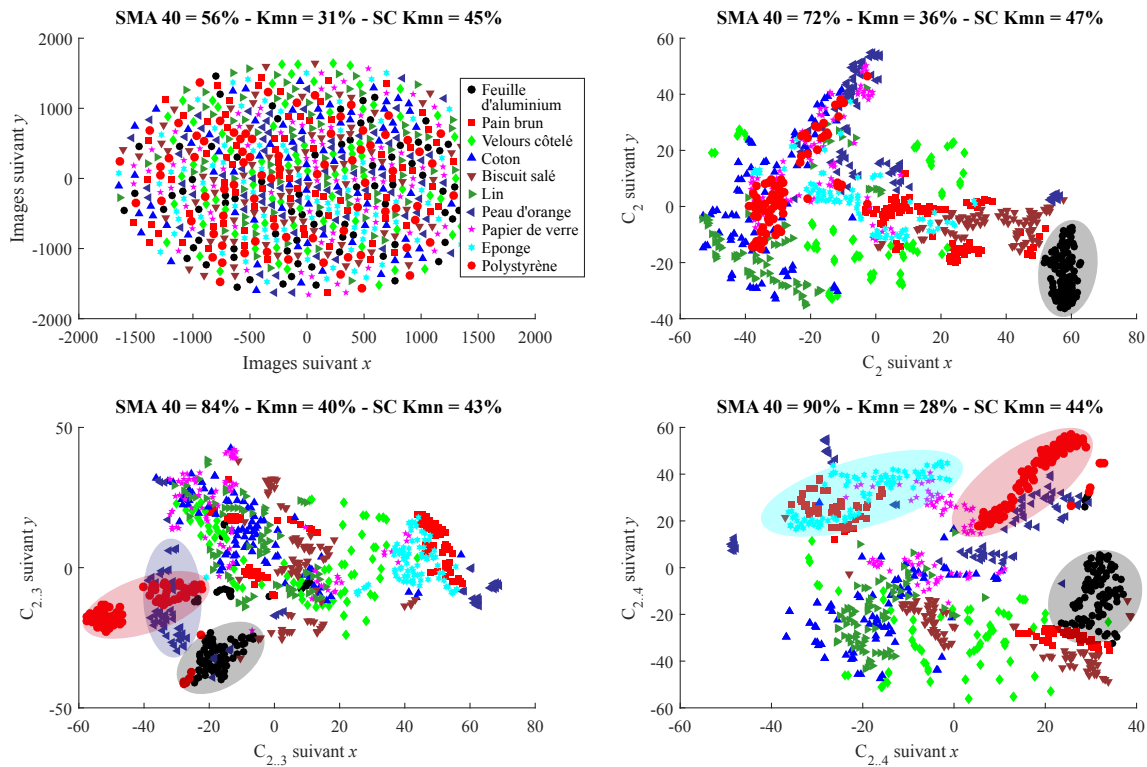


FIGURE 5.9 – *Manifold untangling* des multicorrélations en fonction de l'ordre par la méthode t-SNE : les descripteurs statistiques sont projetés dans un espace à deux dimensions qui préserve les distances (comme à la figure 5.7, les taux de classification sont indiqués au-dessus de chaque sous-figure).

5.1.3.2 Choix des combinaisons d'axes

En sous-section 4.2.2, nous avons évoqué que les multicorrélations d'ordre q sont des tenseurs d'ordre q de taille trop volumineuse pour être utilisées telles quelles pour la classification, voire même pour être simplement calculées. Par conséquent, en sous-section 4.2.3, nous suggérons d'estimer seulement certains axes particuliers (illustré en figure 4.12) issus de la dernière dimension du tenseur d'ordre q . La figure 5.10 donne les résultats de simulation des multicorrélations d'ordre deux. Comme pour la figure 5.8 de la sous-section précédente, nous retrouvons les taux de bonne classification par SMA, avec les deux algorithmes de partitionnement et avec la classification spectrale. Ces taux ont été calculés dans les mêmes conditions que ceux de la figure 5.8 : comparaison avec et sans normalisation par l'écart-type (à la puissance adéquate), trois proportions d'apprentissage pour la SMA, taux moyens sur dix tests sauf pour DBSCAN où est affiché le taux maximal obtenu pour les dix valeurs du nombre minimum de points pour créer un *cluster*.

Globalement, les résultats de la figure 5.10 nous montrent que le taux de bonne classification augmente avec le nombre d'axes considéré. Ce phénomène est particulièrement visible en classification supervisée où les performances sont maximales pour les combinaisons d'axes τ_7 et τ_{11} (gain de 10 % par rapport à τ_1). En classification spectrale, l'ajout d'axes a également permis de gagner quelques points de bonne classification. En revanche, l'apport d'axes supplémentaires est moins flagrant avec les algorithmes des k -moyennes et du DBSCAN mais les taux sont relativement plus faibles (moins de 40 %). Des simulations aux ordres trois et quatre ont donné des résultats analogues, à savoir que l'ajout d'axes supplémentaires dans les descripteurs statistiques permettait d'accroître les performances).

La figure 5.11 illustre l'évolution du taux de bonne classification en fonction de l'augmentation du nombre d'axes (τ_1 , τ_7 et τ_{11}) et du nombre d'ordres des multicorrélations considérés (C_2 , $C_{2,3}$ et $C_{2,3,4}$). Nous pouvons remarquer que l'ajout de l'ordre trois et des axes τ_7 permet une hausse importante des performances, principalement en supervisé où nous atteignons les 95 % de bonne classification. L'apport des multicorrélations d'ordre quatre améliore légèrement les performances en supervisé mais les dégrade en non supervisé. L'ajout des axes supplémentaires permettant de passer de τ_7 à τ_{11} ne semble pas avoir une grande incidence sur le pouvoir discriminant des descripteurs statistiques puisque leurs taux de bonne classification sont quasiment équivalents.

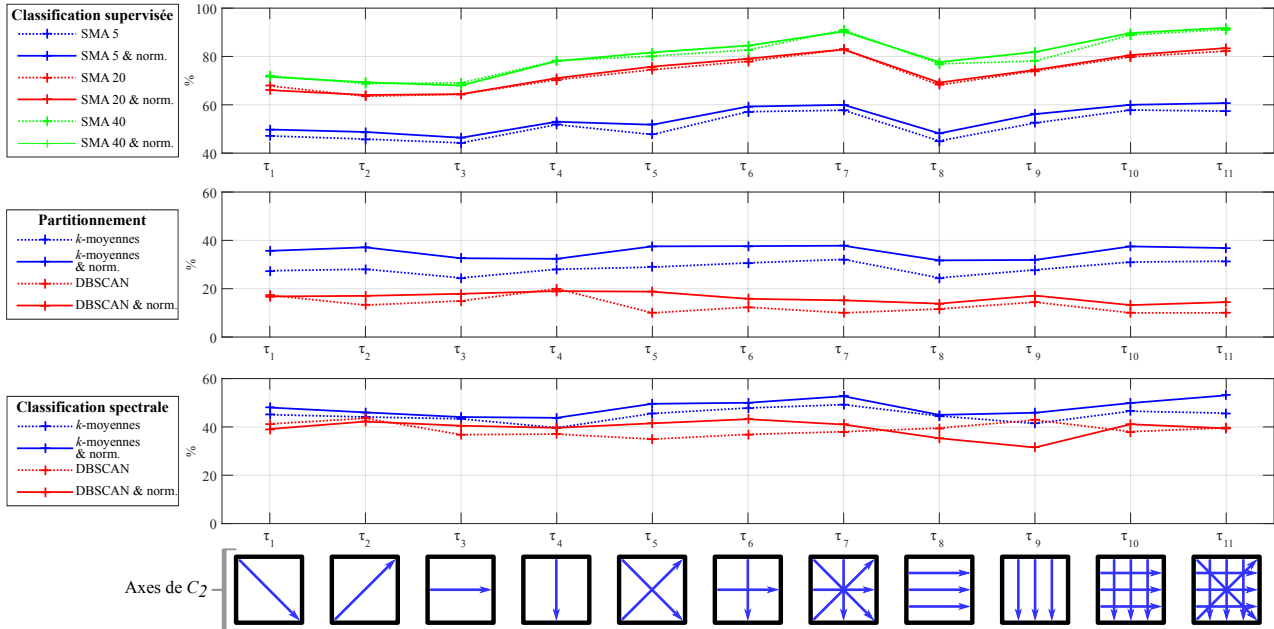


FIGURE 5.10 – Multicorrélations d'ordre deux en fonction des combinaisons d'axes sélectionnés en sous-section 4.2.3 : comme la figure 5.8, cette figure affiche les taux de bonne classification en supervisé, par deux algorithmes de partitionnement et par classification spectrale. La description plus détaillée des axes est donnée en figure 4.12.

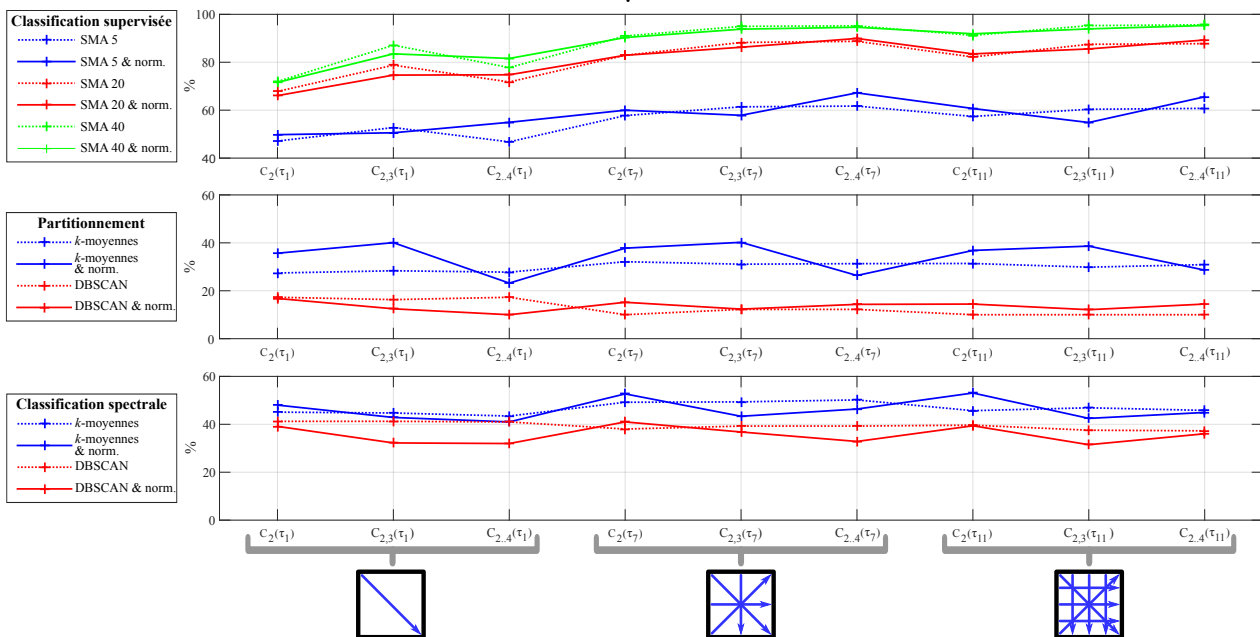


FIGURE 5.11 – Multicorrélations d'ordre supérieur en fonction des combinaisons d'axes et d'ordres suivants : axes τ_1 , τ_7 et τ_{11} en lignes et ordres C_2 , $C_{2,3}$ et $C_{2,3,4}$ en colonnes. Les taux de classification sont calculés de la même manière qu'aux figures 5.8 et 5.10.

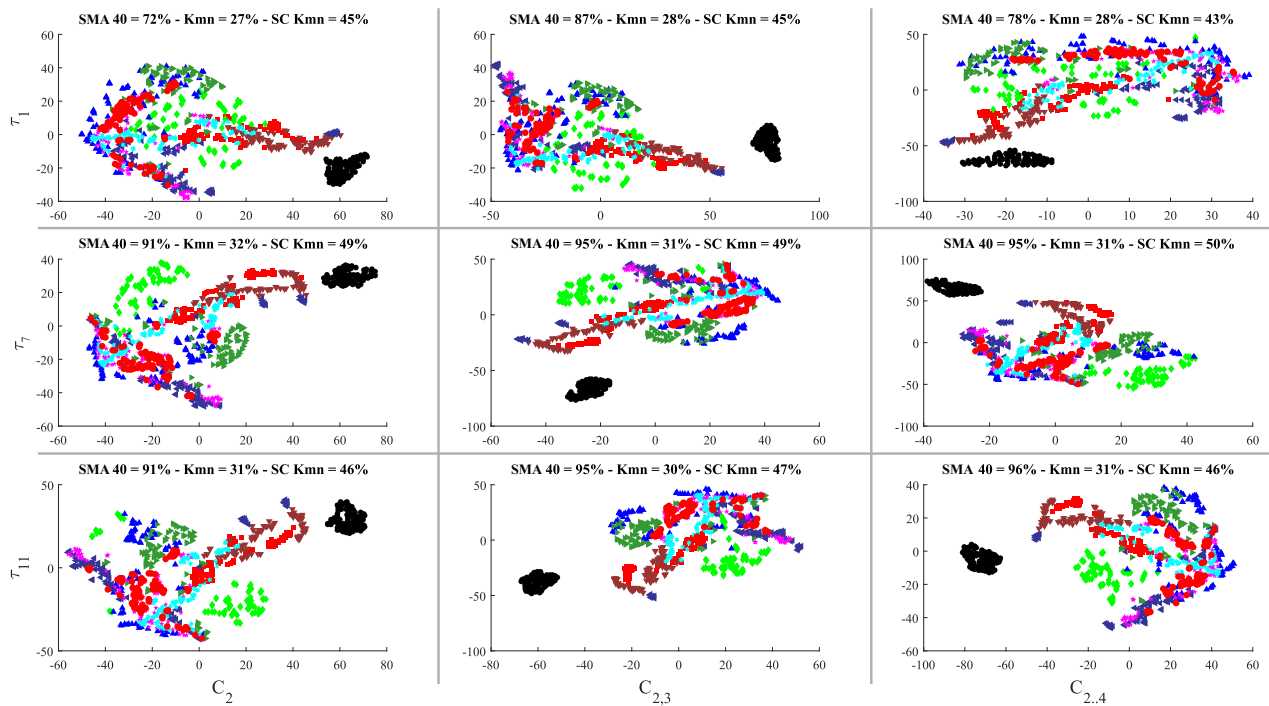


FIGURE 5.12 – *Manifold untangling* des multicorrélations en fonction des combinaisons d’axes et d’ordres par la méthode t-SNE. Les images de la base KTH-TIPS sont représentées par le même symbole et la même couleur qu’à la figure 5.9 en fonction de leur classe.

La figure 5.12 est consacrée, comme l’était la figure 5.9 à la sous-section précédente, au *manifold untangling*. Nous y trouvons la projection 2D des descripteurs statistiques en fonction des combinaisons d’axes et des ordres de la figure 5.11 : τ_1 , τ_7 et τ_{11} en lignes et \mathcal{C}_2 , $\mathcal{C}_{2,3}$ et $\mathcal{C}_{2,3,4}$ en colonnes. En regardant la figure 5.12, deux groupes d’images sautent immédiatement aux yeux car ils se détachent du reste du nuage de points : les classes #1 *Feuille d’aluminium* (ronds noirs) et #3 *Velours* (losanges verts). En scrutant de plus près, les nuages de points semblent être plus disparates en augmentant le nombre d’axes et d’ordres (notamment le nuage central pour τ_7 et $\mathcal{C}_{2,3}$), ce qui confirme que les classes sont plus facilement séparables avec davantage d’axes et d’ordres inclus dans les descripteurs.

5.1.3.3 Bilan sur les descripteurs statistiques

Les simulations exposées dans les deux précédentes sous-sections nous éclairent sur un certain nombre de points quant à l’utilisation des multicorrélations pour résumer des images. Nous avons constaté qu’il est plus judicieux de combiner les multicorrélations calculées pour différents ordres que de s’intéresser à un seul ordre. La normalisation des multicorrélations par l’écart-type à la puissance adéquate (*i.e.* en fonction de leur ordre, *cf.* figure 5.8) donne, en général, des résultats sensiblement meilleurs, principalement lorsque plusieurs ordres de multicorrélations sont considérés. Les ordres deux et trois se sont notamment avérés comme les ordres plus discriminants pour séparer les dix classes d’images alors que les ordres cinq et six ne semblent pas porter d’information supplémentaire qui pourrait améliorer la classification. Ensuite, nous nous sommes intéressés aux calculs des multicorrélations et plus particulièrement aux axes du tenseur que nous allons retenir pour décrire les images. Les résultats ont montré que la combinaison d’axes complémentaires accroît les performances jusqu’à atteindre un seuil auquel l’ajout de nouveaux axes n’améliore plus la classification. Dans cette étude, nous avons sélectionné une liste non exhaustive d’axes élémentaires et, à l’avenir, il serait intéressant d’approfondir les recherches qui permettraient de résumer au mieux les tenseurs des multicorrélations.

Au vu des résultats obtenus dans le cadre de ces simulations, dans la suite de ce mémoire, nous considérerons seulement la combinaison d’axes τ_7 et les multicorrélations seront normalisées par l’écart-type à la puissance adéquate (sauf mention explicite du contraire). En effet, l’ajout d’axes

supplémentaires accroît la taille des descripteurs, et réduit donc la vitesse de calcul des descripteurs et de classification, sans pour autant améliorer de manière significative la classification des images. A présent, si nous comparons le pouvoir discriminant des descripteurs statistiques et multirésolutions, nous constatons que les performances de classification avec le *scattering operator* sont meilleures en supervisé (100 % contre 95 % avec 50 % des images vouées à l'apprentissage) mais quasiment équivalentes en non supervisé (entre 40 et 50 %). Ces résultats ne sont pas forcément illogiques dans le sens où le *scattering operator* est connu pour être performant sur les textures [151] comme celles de la base KTH-TIPS. Dans la prochaine section, nous tenterons de combiner ces deux approches pour améliorer les résultats de classification, notamment ceux en non supervisé.

5.1.4 Descripteurs multirésolutions et statistiques

Dans cette section, nous allons résumer chaque image par un vecteur constitué de ses descripteurs multirésolutions et statistiques, comme évoqué en section 4.3. Ces deux types de descripteurs pourraient s'avérer complémentaire puisque le *scattering operator* s'intéresse aux éléments de textures et les statistiques d'ordre supérieur à la distribution des pixels dans les images. En sous-section 5.1.4.1, nous nous interrogerons sur l'uniformisation des descripteurs afin d'exploiter au mieux leurs propriétés discriminantes. La sous-section 5.1.4.2 compare les performances des approches seules et combinées pour différents jeux de paramètres retenus aux sous-sections 5.1.3.1 et 5.1.3.2.

5.1.4.1 Uniformisation des descripteurs

Contrairement aux sections précédentes où le SO et les SOS sont appliqués séparément, cette section propose d'associer ces deux approches. Les descripteurs étant de nature différente, nous avons suggéré, en section 4.3.2, huit manières d'uniformiser les deux types de descripteurs obtenus avant l'étape de classification. Celles-ci sont rappelées dans le tableau 5.4 où sont distinguées les techniques d'uniformisation et les manières de les appliquer aux descripteurs :

Techniques	Applications
1. Standardisation (centrage et réduction)	A. Par type sur l'ensemble des images
2. Normalisation <i>min-max</i>	B. Par ordre sur l'ensemble des images
	C. Par type et par image
	D. Par ordre et par image

Tableau 5.4 – Méthodes d'uniformisation des descripteurs multirésolutions et statistiques.

Les résultats de classification en fonction des uniformisations sont donnés en figure 5.13. Les descripteurs multirésolutions ont été obtenus par l'application classique du SO (*i.e.* invariance en translation seule) avec ou sans la linéarisation des ordres par le logarithme (respectivement colonne de droite et de gauche sur la figure 5.13). Les descripteurs statistiques correspondent aux multicorrélations d'ordres deux, trois et quatre calculées suivant la combinaison d'axes τ_7 (vertical, horizontal et diagonal). Si nous regardons les trois premiers points (quelque soit le mode de classification et l'usage du logarithme), nous constatons que les performances des SOS sont légèrement en-dessous de celles du SO, qui sont quasiment équivalentes à l'utilisation combinée des deux approches sans uniformisation. En comparant globalement les courbes de gauche (*SO* sans logarithme) et celles de droite (*SO* avec logarithme), nous retrouvons le constat exposé à la section 5.1.1 sur l'apport du logarithme pour la classification des descripteurs du SO.

En uniformisant l'ensemble des descripteurs par type d'approche (les deux points suivants correspondant à la lettre A), nous observons des performances proches de celles obtenues sans uniformisation (Σ_{brute}). Les performances deviennent plus intéressantes lorsque l'uniformisation est réalisée par ordre (du SO et des SOS) sur l'ensemble des images (les deux points suivants correspondant à la lettre B). Ces taux de bonne classification sont en effet légèrement supérieurs à ceux obtenus avec les autres manières d'uniformiser les descripteurs A, C et D, particulièrement en classification spectrale (jusqu'à 10 % de gain). L'uniformisation par image, et notamment par type

(baisse de quasiment 10 % des taux), ne semble pas être un choix judicieux pour classer les images de la base KTH-TIPS. Cependant, sur les images de thrombi (ou de glandes), elle pourrait s'avérer intéressante pour limiter l'impact des changements de paramètres d'acquisition entre les patients.

En ce qui concerne la technique d'uniformisation (1 ou 2), les résultats sont plutôt similaires en supervisé et plus disparates en non supervisé. De manière générale, la normalisation semble être un choix cohérent par rapport à l'ensemble des résultats. Des simulations menées avec d'autres paramètres des multicorrélations (*i.e.* d'autres combinaisons d'ordre et d'axes) ont donné des résultats analogues (non présents dans ce mémoire).

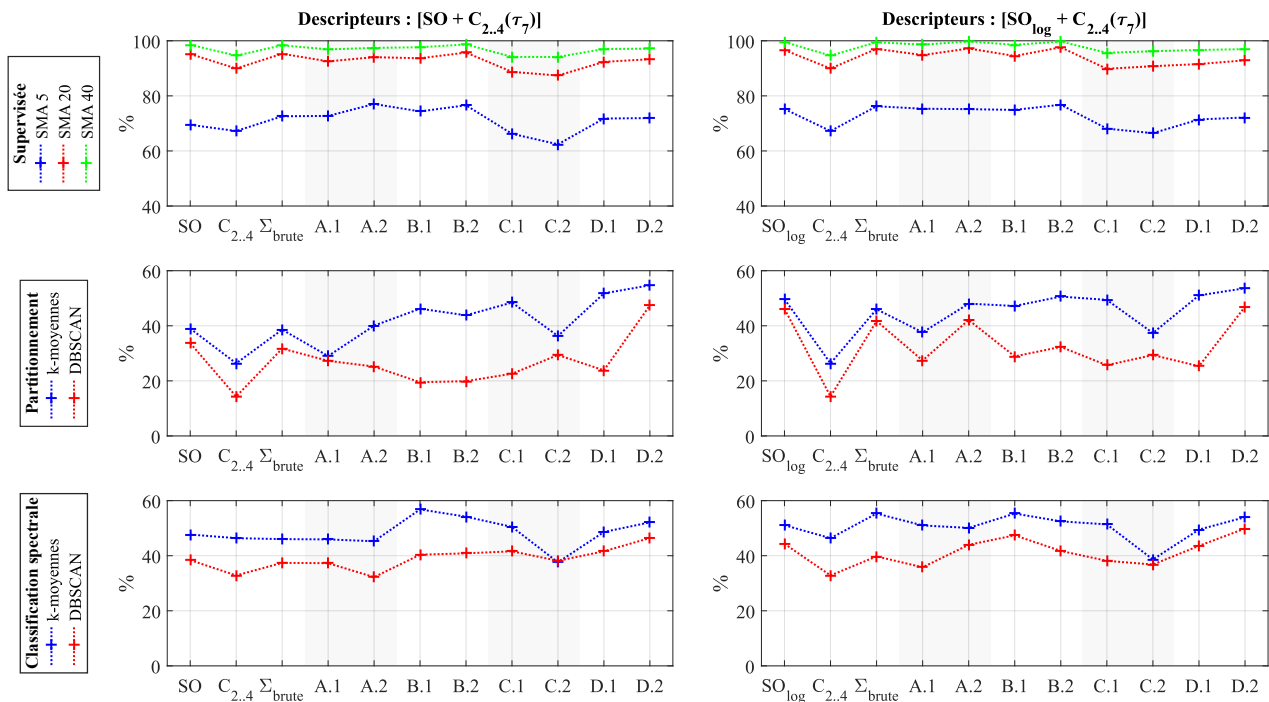


FIGURE 5.13 – Uniformisation des descripteurs multirésolutions et statistiques : chaque graphe donne les taux de bonnes classifications pour les descripteurs SO et SOS seuls et associés sans (Σ_{brute}) ou avec chacune des huit combinaisons d'uniformisation (de A.1 à D.2). Comme à la figure 5.8, les lignes correspondent aux modes de classification (mode supervisé, partitionnement et classification spectrale). Les graphes de la colonne de droite (respectivement de gauche) ont été obtenus sans (respectivement avec) l'utilisation du logarithme pour linéariser les coefficients du SO par ordre.

5.1.4.2 Evaluation des performances

Les simulations de la sous-section précédente nous ont permis d'établir la méthode d'uniformisation des descripteurs multirésolutions et statistiques qui sera utilisée dans les prochaines sections : la normalisation par ordre sur l'ensemble des descripteurs. Dans cette sous-section, nous allons vérifier l'intérêt de combiner le SO et les SOS en appliquant cette méthode d'uniformisation pour différents ordres des multicorrélations : $C_{1..4}$, C_2 , $C_{2,3}$ et $C_{2..4}$. Pour ce faire, nous allons nous appuyer sur la figure 5.14 qui donne les taux de bonne classification par SMA (première ligne), par partitionnement (deuxième ligne) et par classification spectrale (troisième ligne).

En analysant la figure 5.14, nous pouvons voir des courbes en forme de vagues (principalement en supervisé) où le creux correspondant à l'utilisation seule des SOS et la crête à celle des deux types de descripteurs uniformisés. Par exemple, avec seulement 6.25 % des images consacrées à l'apprentissage, les descripteurs mixtes (Σ_{npo}) permettent de gagner, respectivement, 10 % et 20 % par rapport aux descripteurs multirésolutions et statistiques seuls. De même, en classification spectrale, les deux types de descripteurs uniformisés permettent d'atteindre quasiment 60 % de bonne classification alors que les performances se situent davantage au niveau des 50 % en les considérant de manière indépendante.

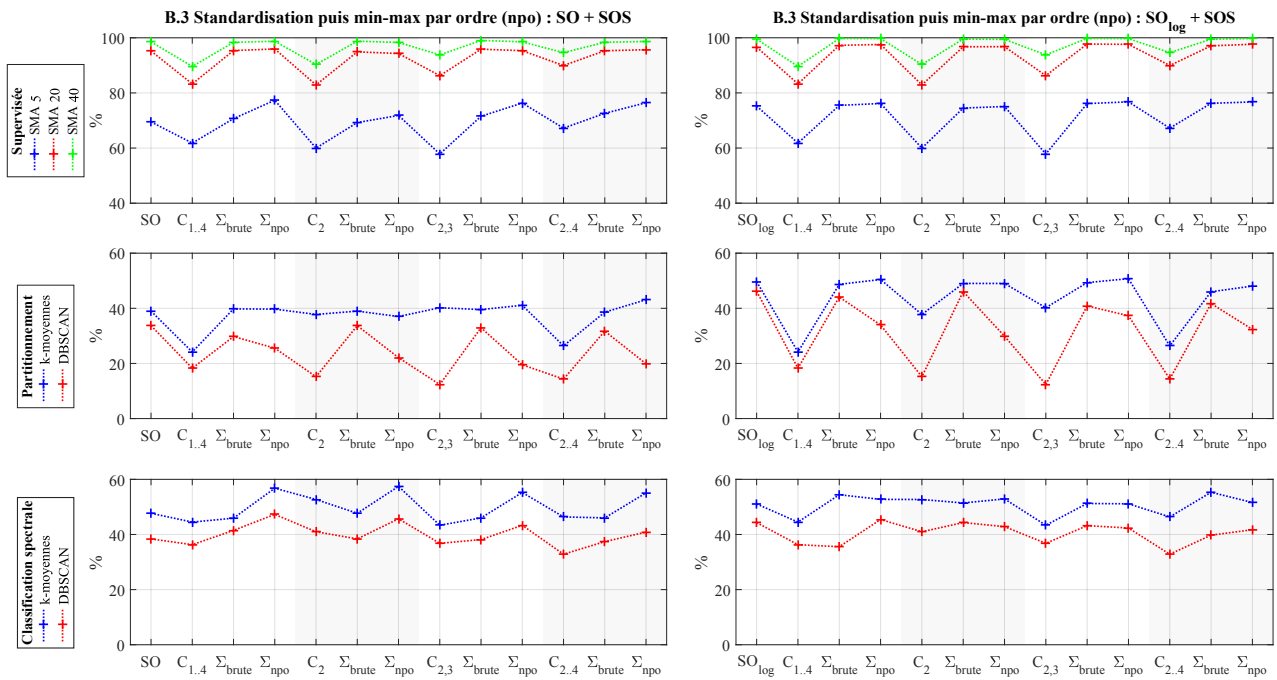


FIGURE 5.14 – Classification des descripteurs multirésolutions et statistiques uniformisés par la normalisation *min-max* des ordres de l'ensemble des descripteurs : chaque graphe donne les taux de bonnes classifications pour les descripteurs SO et SOS ($C_{1.4}(\tau_1)$, $C_2(\tau_7)$, $C_{2,3}(\tau_7)$ ou $C_{2.4}(\tau_7)$) seuls et associés sans (Σ_{brute}) ou avec uniformisation (Σ_{npo}).

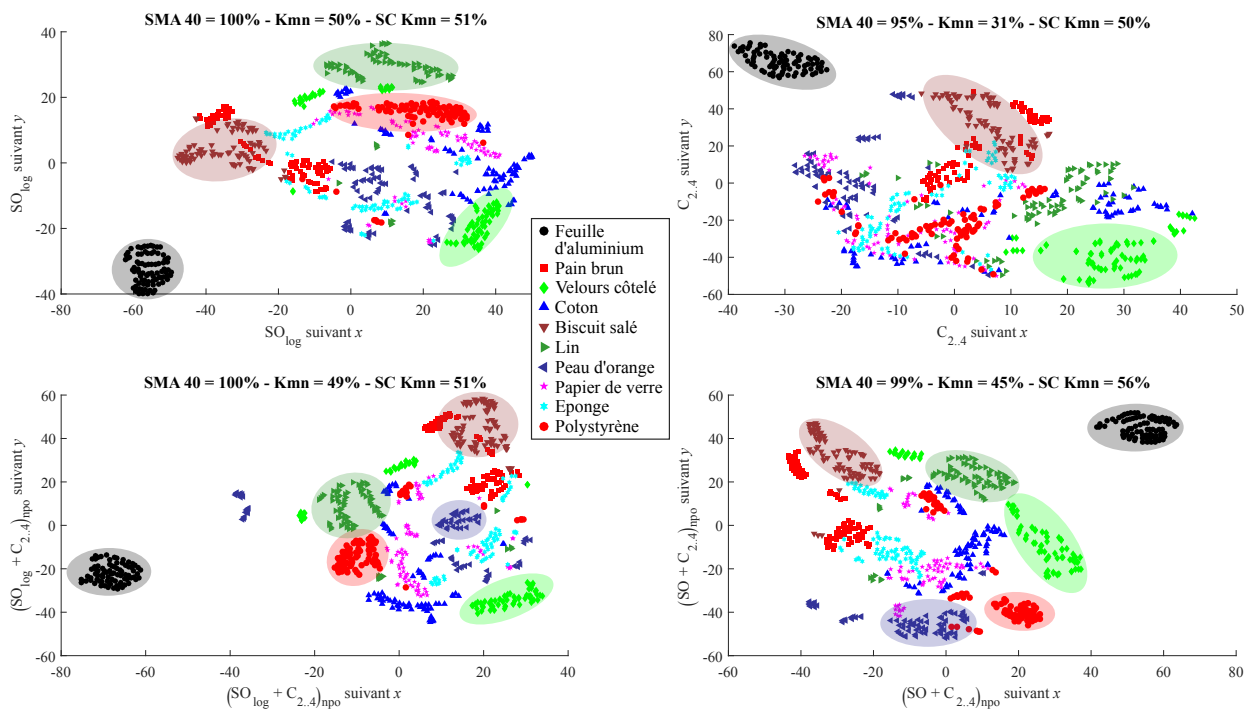


FIGURE 5.15 – *Manifold untangling* des descripteurs multirésolutions et statistiques (voir figure 5.7 pour plus d'information).

5.1.4.3 Bilan sur l'utilisation conjointe du SO et SOS

Dans cette section, nous nous sommes penchés sur la complémentarité des approches multirésolutions et statistiques. Les résultats présentés dans les deux sous-sections précédentes montrent que leur utilisation conjointe est pertinente puisqu'elle permet d'améliorer les performances de classification. Pour réunir ces deux types de descripteurs différents, la normalisation *min-max* par ordre sur l'ensemble des descripteurs paraît être une solution adaptée en donnant à chaque ordre des descripteurs un poids similaire lors de la classification. En effet, l'absence d'uniformisation donne davantage de poids aux ordres inférieurs, qui sont plus énergétiques. Du point de vue du *manifold untangling* illustré en figure 5.15, nous pouvons apercevoir l'évolution positive entre l'espace initial en 2D (voir figure 5.9), l'espace en 2D des SOS, celui du SO et l'espace des descripteurs combinés (SO et SOS). Dans ce dernier, les classes *Feuille d'aluminium* (#1), *Velours* (#3), *Lin* (#6), *Peau d'orange* (#7) et *Polystyrène* (#10) semblent plus facilement séparables que précédemment.

5.1.5 Bilan des simulations sur la base de référence

Cette section s'est articulée autour d'une base de référence (images synthétiques de textures) dans le but de vérifier l'implémentation des approches choisies pour caractériser les thrombi et de prédéfinir leur paramétrage afin de maximiser leur chance de succès sur nos données réelles. En consacrant 25 % des images à l'apprentissage, la description des images par le *scattering operator*, les statistiques d'ordre supérieur ou les deux approches donnent de très bons résultats avec plus de 90 % de bonne classification. Avec 6.25 % des images pour l'apprentissage (*i.e.* cinq images par classe en supervisé), les performances oscillent entre 60 et 80 % : les meilleurs taux étant obtenus pour le *scattering operator* avec le logarithme (et l'apprentissage multi-échelle), et l'utilisation combinée des deux approches. Les performances en non supervisé, entre 40 et 60%, sont inférieures à celles obtenues en supervisé mais restent acceptables sachant qu'il y a dix classes dans cette base de référence.

Globalement, sur cette base constituée d'images texturées, le *scattering operator* paraît plus adapté que les multicorrélations mais celles-ci, associées aux descripteurs multirésolutions, permettent tout de même d'améliorer les performances de classification, notamment en non supervisé. De plus, les statistiques d'ordre supérieur offrent des perspectives d'amélioration comme la contraction de tenseur ou l'étude des multispectres (*cf.* section 4.2.3). La base de thrombi étant encore de taille restreinte, l'augmentation de données (comme l'apprentissage multi-échelle) pourrait se montrer intéressante pour doper la taille de la base en classification non supervisée. Enfin, en observant les projections 2D des descripteurs, nous avons pu constater que l'utilisation de ces approches a bien permis de simplifier la séparation des classes bien que certaines soient encore mélangées dans ce plan (ce qui ne présage pas qu'elles ne soient pas séparables dans l'espace initial, avant la projection). L'astuce du noyau, qui permet d'identifier des frontières non linéaires (*cf.* sous-section 2.4.1.2), pourrait également s'avérer comme un axe d'amélioration prometteur pour la classification des images, notamment en mode non supervisé.

5.2 Evaluation sur les glandes salivaires

Précédemment, nous avons validé l'implémentation et paramétré les approches de caractérisation (multirésolution et statistique) et de classification sur une base de référence composée d'images synthétiques. Avant de s'attaquer aux échographies et aux élastographies des thrombi, cette section envisage de tester nos approches sur une base de même nature mais plus simple que celle des thrombi : les glandes salivaires. L'objectif des simulations est alors de classer les glandes en deux catégories, normales ou pathologiques. Contrairement à notre problématique sur les thrombi, l'échographie des glandes salivaires permet visuellement à un expert médical d'écarter ou de suspecter le syndrome de Gougerot-Sjörden. De plus, le comportement des ondes de cisaillement dans les glandes, qui sont du tissu, est plus conforme à la théorie utilisée pour générer les élastographies que dans les thrombi et leur structure gélatineuse (*cf.* section 1.3.3). La sous-section 5.2.1 présente un bref état de l'art sur le diagnostic du syndrome de Gougerot-Sjörden par échographie et élastographie. Ensuite, la sous-section 5.2.2 expose le protocole expérimental et la sous-section 5.2.3 résume les résultats des simulations menées sur la base des glandes salivaires.

5.2.1 Diagnostic des glandes salivaires par échographie et élastographie

En médecine, deux critères statistiques sont généralement utilisés pour évaluer une approche de diagnostic : la sensibilité et la spécificité.

La *sensibilité* d'un diagnostic est sa capacité à détecter les patients atteints par la maladie en question, le syndrome de Gougerot-Sjörden dans notre étude [139]. Elle est définie en fonction du nombre (ou du pourcentage) de Vrais Positifs (VP) et de Faux Négatifs (FN) par la relation :

$$\text{Sensibilité} = \frac{\text{VP}}{\text{VP} + \text{FN}}. \quad (5.1)$$

La *spécificité* d'un diagnostic correspond au taux de détection des patients non atteints par la maladie en question. Elle est estimée à partir du nombre de Vrais Négatifs (VN) et de Faux Positifs (FP) par :

$$\text{Spécificité} = \frac{\text{VN}}{\text{VN} + \text{FP}}. \quad (5.2)$$

A titre de référence, les auteurs de [80] ont réalisé une évaluation comparative d'études sur la détection (par des experts médicaux) du syndrome de Gougerot-Sjödren par échographie. Les onze études évoquées dans cette publication (entre 1996 et 2012) affichent une sensibilité moyenne de 77.9 % (allant de 46.6 à 90 %) et une spécificité moyenne de 86.7 % (allant de 73 à 100 %). Dans une autre publication [41], les auteurs s'intéressent à des études plus récentes (de 2010 à 2015) qui présentent une sensibilité moyenne de 71.3 % (de 63 à 91 %) et une spécificité moyenne de 91.9 % (83 à 98 %). Lors de ces études, les glandes parotides et submandibulaires sont mélangées.

Des articles récents se sont penchés sur l'intérêt de l'élastographie dans le diagnostic de ce syndrome et ont montré que l'élastométrie (valeurs moyennes de l'élastographie dans une région d'intérêt) des glandes parotides pathologiques donne des valeurs plus élevées (ce qui indique une glande plus dure) que les glandes normales [83, 140]. L'étude présentée en [83] a montré que l'élastométrie des glandes parotides permettait de distinguer les patients sains des patients atteints du syndrome avec une sensibilité de 81 % et une spécificité de 67 %. En revanche, sur les glandes submandibulaires, l'élastométrie n'a pas permis, dans les deux études [83, 140], de séparer les deux groupes de patients.

5.2.2 Protocole expérimental

Le protocole expérimental des simulations présentées à la sous-section 5.2.3 s'articule en quatre parties : un rappel sur la base, la définition des descripteurs, leur uniformisation et leur classification.

5.2.2.1 Rappel sur la base

La base de données des glandes salivaires, introduite en section 1.3.3, est constituée, comme la base de thrombi veineux, d'images échographiques et élastographiques collectées avec le système Toshiba Aplio 500. Ces images ont été récoltées auprès de trente-sept patients, ce qui donne un total d'environ cent quarante glandes parotides (droite et gauche) et submandibulaires (soit soixante-dix de chaque). Le but de l'examen médical est de détecter, chez un patient qui souffre d'une sécheresse buccale et/ou oculaire, le syndrome de Gougerot-Sjörden. Ce syndrome est une maladie auto-immune qui se traduit par une texture des glandes plus hétérogène et plus échogène sur les échographies. Généralement, les glandes atteintes de cette maladie sont également plus dures, notamment pour les parotides, ce qui justifie l'intérêt de l'élastographie. Bien qu'il existe trois grades de la maladie en fonction de sa gravité, notre étude est simplifiée au classement binaire des glandes, normales ou Gougerot-Sjörden, car la grande majorité des glandes pathologiques sont du grade trois (le plus avancé).

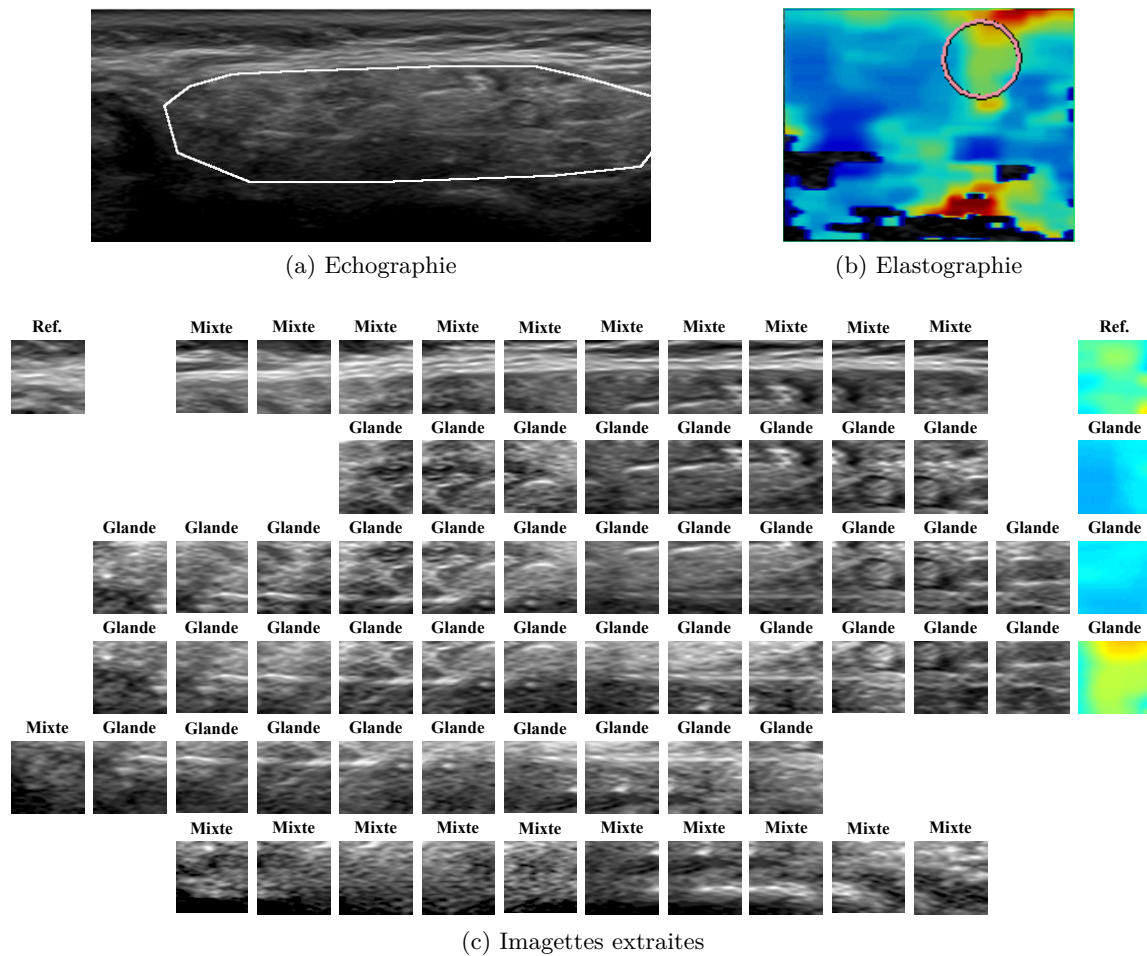


FIGURE 5.16 – Imagettes d'une glande parotide malade tirées de l'échographie (a), de l'élastographie (b) et étiquetée en trois catégories (voir les sections 3.3.2 et 3.3.3 pour plus d'information) : « Glande », « Mixte » et « Référence ».

La base de données est constituée à soixante pour cent de glandes saines et quarante pour cent de glandes pathologiques. Les préparations des échographies et élastographies font l'objet des sections 3.3.2 et 3.3.3. Au total, des centaines et des dizaines d'imagettes ont été extraites, respectivement, des échographies et des élastographies. La figure 5.16 propose un exemple d'imagettes récupérées à partir d'une acquisition de glande parotide pathologique. Ces imagettes sont étiquetées en trois catégories (deux pour les élastographies, *cf.* section 3.3.3) en fonction du pourcentage de glande qu'elle contient : « Glande », « Mixte » et « Référence ». Afin de prendre en compte le contraste entre la glande et les structures avoisinantes, nous proposons de standardiser puis de normaliser (*cf.* section 4.3.2) chacune des imagettes par rapport à l'ensemble des imagettes de l'acquisition. Cela signifie qu'une imagette est

standardisée par la valeur moyenne et l'écart-type de toutes les imgettes, puis, de la même manière, normalisée par leurs valeurs minimale et maximale. Cette uniformisation peut également présenter l'avantage de limiter l'impact des changements de paramètres d'acquisition entre patients. Dans le même but, les histogrammes des échographies ont été égalisés avant l'extraction des imgettes par la méthode CLAHE avec une distribution de Rayleigh³ (cf. section 3.2.3). Les simulations exposées par la suite compareront donc les performances de classification avec et sans cette uniformisation, avec et sans égalisation d'histogramme.

5.2.2.2 Définition des descripteurs

Chaque imgette de la base de données (« Mixtes » et « Références » comprises) est représentée par des descripteurs multirésolutions et/ou statistiques. Ces derniers ont été calculés pour différents jeux de paramètres définis dans les deux paragraphes suivants.

Le SO est appliqué jusqu'à l'ordre deux, avec huit orientations et un nombre d'échelles qui dépend de la taille des imgettes N :

$$J = \log_2 \left(\sqrt{N} \right) - 1 \tag{5.3}$$

où N vaut 64×64 pour les échographies, 32×32 pour les élastographies. Ce choix de J , préconisé par les auteurs qui ont intronisé le SO [22], a été validé sur la base de référence KTH-TIPS où les simulations, exposées en sous-section 5.1.2.1, montrent qu'il est cohérent. Lors des simulations, nous comparerons les performances du SO avec ou sans l'utilisation du logarithme pour linéariser les coefficients par ordre (cf. sous-section 5.1.2.2).

Les multicorrélations sont calculées de l'ordre un à l'ordre quatre, pour la combinaison d'axes τ_7 (horizontal, vertical et diagonaux) et chaque ordre est normalisé par l'écart-type à la puissance adéquate. Durant les simulations, nous comparerons les performances de classification pour différentes associations d'ordres des multicorrélations : $\mathcal{C}_2(\tau_7)$, $\mathcal{C}_{1,2}(\tau_7)$, $\mathcal{C}_{2..4}(\tau_7)$ et $\mathcal{C}_{1..4}(\tau_7)$. Ensuite, les descripteurs du SO et ces cumulants seront combinés pour vérifier si cela permet d'améliorer la classification des images, comme nous l'avons constaté sur la base de référence.

5.2.2.3 Uniformisation des descripteurs

Au final, lorsque toutes les imgettes de la base sont décrites par le SO et les SOS, nous obtenons les vecteurs descripteurs tels qu'illustrés par la figure 5.17. Chaque glande est donc représentée par autant de vecteurs descripteurs qu'il y a d'imgettes extraites de ses acquisitions. A la sous-section 5.2.2.1, nous avons suggéré de normaliser les imgettes de chaque acquisition pour prendre en compte les variations de contraste entre la glande et son voisinage mais aussi pour limiter l'influence des variations des paramètres d'acquisition. Dans le même objectif, les descripteurs des imgettes seront normalisés pour chaque acquisition. Ce procédé est appelé *uniformisation des descripteurs par acquisition* dans la suite de ce mémoire. Les performances obtenues avec ce procédé seront comparées à celles obtenues sans uniformisation. De plus, nous avons vu en sous-section 5.1.4 que l'uniformisation par ordre sur l'ensemble des données, appelée *uniformisation globale par ordre*, permettait également d'augmenter les performances, notamment lorsque les deux types de descripteurs sont combinés.

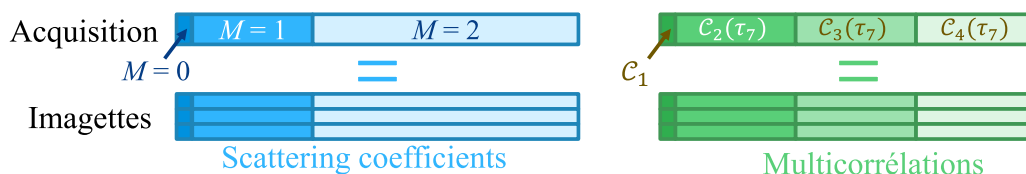


FIGURE 5.17 – Descripteurs multirésolutions et statistiques des imgettes : chaque acquisition est représentée par les vecteurs descripteurs de ces imgettes de « Glande » (cf. figure 5.16).

3. Les simulations donnent des résultats analogues avec les autres méthodes proposées.

5.2.2.4 Fusion des descripteurs

Actuellement, chaque glande est représentée par les vecteurs descripteurs calculés sur chacune des imagerie extraites de ses acquisitions. Pour accroître la stabilité de la représentation des glandes, nous suggérons de moyenner les vecteurs descripteurs obtenus sur une même acquisition. Le prix de cette stabilité accrue des descripteurs peut-être une perte de détails qui pourrait nuire à leur pouvoir discriminant. Par exemple, si une glande parotide a été échographiée à deux reprises et que vingt imagerie de glande ont été extraites de ces deux acquisitions, alors, après fusion des descripteurs, cette glande sera décrite par deux vecteurs, correspondant à la moyenne des vecteurs descripteurs des imagerie de chaque acquisition.

5.2.2.5 Classification des descripteurs

Comme pour la base de référence, les descripteurs des glandes salivaires sont classés de manière supervisée par SMA et non supervisée par k -moyennes, DBSCAN et la Classification Spectrale (CS). Chaque glande sera considérée comme pathologique lorsqu’au moins 50 % de ces imagerie (ou de ses acquisitions s’il y a fusion des descripteurs) seront classées comme *Gougerot* (pour Gougerot-Sjörden).

En supervisé, trois proportions apprentissage/test ont été choisies pour les simulations : 50 %, 25 % ou 12.5 % des glandes de chaque classe (*Normale* ou *Gougerot*) sont consacrées à l’apprentissage. Pour chaque proportion, les taux de classification sont calculés pour dix partitions aléatoires différentes puis moyennés. Par SMA, une imagerie de test donnée est attribuée à la classe dont le modèle affine est le plus proche de son vecteur descripteur. Par conséquent, nous proposons de rattacher la glande à la classe de l’imagerie qui sera la plus proche (parmi toutes les imagerie de la glande) d’un des deux modèles de classe. Les taux obtenus avec cette règle de décision, notée \mathcal{R}_2 seront comparés avec ceux obtenus avec la règle émise précédemment (classe la plus étiquetée parmi les imagerie), notée \mathcal{R}_1 .

En classification non supervisée, les performances sont estimées de la même manière que sur la base KTH-TIPS (*cf.* section 5.1.3.1) pour les quatre techniques de classification utilisées. L’algorithme des k -moyennes, appliqué directement sur les descripteurs ou pendant la CS, est appliqué pour deux *clusters* et pour dix initialisations aléatoires des deux *clusters* (*i.e.* dix applications de l’algorithme). DBSCAN, sur les descripteurs ou lors de la CS, est appliqué pour les dix mêmes valeurs du paramètre v qui correspond au nombre minimum de points pour définir un *cluster* : 3, 5, 7, 8, 9, 10, 11, 12, 15 et 20. Le taux de classification correspond au meilleur des taux obtenus pour chacun de ces dix paramètres.

Enfin, pour la CS, le paramètre de dispersion, qui permet de calculer la similarité gaussienne entre les descripteurs de la base, a été défini, comme pour la base de référence, en fonction du voisinage de chaque point. Pour déterminer le nombre de voisins r à considérer, les simulations présentées en annexe E ont montré que le rapport entre le nombre total d’images de la base de données et le nombre de classes est un choix pertinent de r . Lors des simulations réalisées sur la base de glandes salivaires, le nombre de voisins r est donc défini par :

$$r = \frac{\text{nombre total de glandes}}{\text{nombre de classes}} \quad (5.4)$$

Taux (%)	100	97.5	95	92.5	90	87.5	85	82.5	80	77.5	75	72.5	70	67.5	65	62.5	60	57.5	55	52.5	50	
Couleur																						

(a) Classification supervisée

Taux (%)	80	77.5	75	72.5	70	67.5	65	62.5	60	57.5	55	52.5	50	47.5	45	42.5	40	37.5	35	32.5	30	
Couleur																						

(b) Classification non supervisée

Tableau 5.5 – Echelle des couleurs en fonction du taux de bonne classification.

Les résultats de classification pour chaque simulation (lignes) et jeux de paramètres (colonnes) seront donnés sous forme de tableaux où chaque case indique le taux de bonne classification moyen (ou maximal pour le DBSCAN) en pourcentage. Chaque case est colorée en fonction du pourcentage suivant les échelles des couleurs illustrées dans le tableau 5.5. La classification supervisée, donnant des résultats plus élevées, pour la base KTH-TIPS, que la non supervisée, les couleurs de leurs échelles ont donc été échelonnées de manières différentes.

5.2.2.6 Résumé du protocole

L'ensemble du protocole expérimental, utilisé pour obtenir les résultats présentés à la section 5.2.3, est résumé sur la figure 5.18. Notre base est constituée de trois types d'images : les échographies avec ou sans égalisation d'histogramme, et les élastographies. Pour limiter l'influence des réglages des modes d'imagerie et prendre en compte les structures voisines de la glande, les imagerie de chaque glande pourront être uniformisées par rapport à l'ensemble des imagerie extraites d'une même acquisition (glandes, mixtes ou référence). Dans le même objectif, cette uniformisation pourra également avoir lieu sur les descripteurs des imagerie (au lieu ou en plus de l'uniformisation des imagerie). Les imagerie sont représentées par des descripteurs multirésolutions et/ou statistiques calculés avec les paramètres affichés sur la figure 5.18.

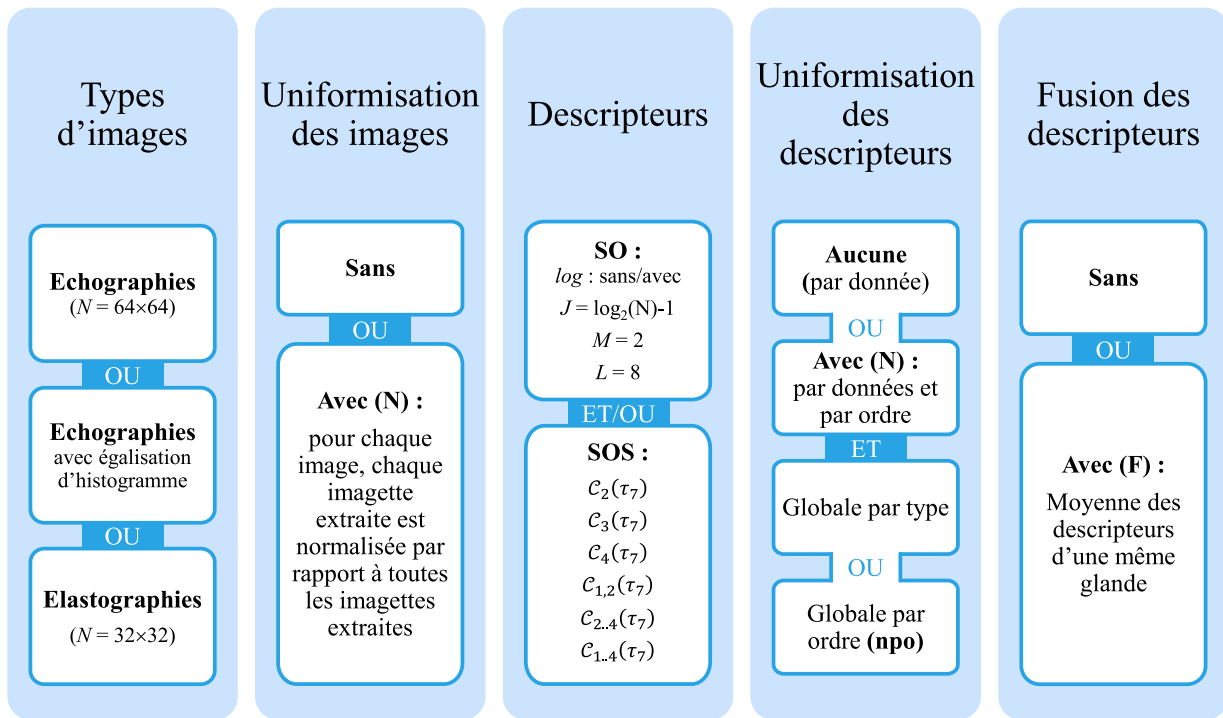


FIGURE 5.18 – Protocole expérimental des simulations sur les glandes salivaires.

En plus des deux uniformisations évoquées précédemment, l'ensemble des descripteurs sera uniformisé globalement par type (SO et SOS) ou par ordre. En effet, sur la base de référence, nous avons constaté que ce procédé permettait d'accroître les performances de classification avec les descripteurs combinés (SO et SOS). La fusion des descripteurs des imagerie d'une même acquisition devrait permettre d'accroître la stabilité de la description des acquisitions. A travers les simulations, nous observerons si la fusion des descripteurs par acquisition permet effectivement d'améliorer la classification des glandes ou si elle engendre une perte d'information qui nuit à la bonne classification des descripteurs. Au final, nous obtenons deux fois vingt-quatre (les uniformisations globales par type ou par ordre ne seront pas affichées dans les mêmes tableaux) combinaisons différentes de ces paramètres de simulation qui seront appliqués sur dix-sept jeux de descripteurs différents : images vectorisées, deux paramétrages du SO, six définitions des multicorrélations et la combinaison de deux paramétrages du SO avec quatre de six définitions des multicorrélations.

5.2.3 Résultats du protocole

Cette section présente seulement la partie la plus représentative des résultats car l'ensemble des simulations réalisées dans le cadre du protocole expérimental décrit à la section précédente nécessiterait la publication de soixante-quinze tableaux du type de ceux qui sont présentés dans cette section. La sous-section 5.2.3.1 considère les glandes parotides et submandibulaires communément alors que les sous-sections 5.2.3.2 et 5.2.3.3 les traitent de manière indépendante.

5.2.3.1 Parotides et submandibulaires

Dans cette sous-section, nous ne faisons pas la distinction entre les glandes parotides et les glandes submandibulaires. En supervisé par exemple, l'apprentissage des glandes pathologiques se fera aussi bien sur des vecteurs descripteurs de glandes parotides que submandibulaires. Nous allons d'abord nous intéresser aux résultats obtenus par SMA avec 50 % des images en apprentissage puis nous réduirons progressivement la proportion d'images dans la base d'apprentissage jusqu'à atteindre la classification sans apprentissage. En non supervisé, nous analyserons les performances des deux algorithmes de partitionnement (k -moyennes et DBSCAN) et de la Classification Spectrale (CS)).

1. Classification supervisée par Sélection de Modèles Affines (SMA) :

Nous allons tout d'abord nous pencher sur les résultats de la classification par SMA avec 50 % des glandes dédiées à l'apprentissage. Les pourcentages de bonne classification sont affichés dans le tableau 5.6. Globalement, nous pouvons observer que les taux obtenus à partir des échographies (vers les 70 % en moyenne) sont supérieurs à ceux des élastographies (vers les 60 %). Au niveau des échographies, les simulations montrent de manière flagrante l'intérêt d'utiliser des techniques d'extractions de descripteurs plutôt que de classer directement les échographies vectorisées (cases rouges aux alentours des 40 % de classification). En revanche, ce n'est pas le cas pour les élastographies pour lesquelles nos deux approches de description semblent aussi peu performantes que les images vectorisées (voire moins pour les multicorrélations seules).

En comparaison des règles de décision \mathcal{R}_1 (classe majoritaire parmi les imagerie) et \mathcal{R}_2 (classe de l'imagerie la plus proche d'un des deux modèles), les tableaux 5.6.a et 5.6.b correspondants montrent des résultats assez similaires bien que des derniers semblent légèrement meilleurs avec la règle \mathcal{R}_2 . Par la suite, l'ensemble des résultats en classification supervisée par SMA sont donnés avec cette règle \mathcal{R}_2 . Le tableau 5.6.c présente les résultats acquis en uniformisant globalement les descripteurs par ordre (et avec la règle \mathcal{R}_2). D'une manière générale, et dans ce cas particulier de la SMA avec 50 % de glandes pour l'apprentissage, les performances de classification, en uniformisant les descripteurs de la sorte, sont légèrement meilleures que celles obtenues en uniformisant uniquement par type de descripteurs (SO et SOS). Les prochains tableaux de résultats de cette section sur les glandes salivaires incluront cette uniformisation par ordre.

En se focalisant sur ce tableau 5.6.c, nous remarquons que les meilleures performances sont obtenues en combinant les descripteurs multirésolutions (SO_{\log}) et statistiques (\mathcal{C}_2 ou $\mathcal{C}_{1,2}$) avec un pourcentage pouvant atteindre les 80 % de bonne classification. Le SO seul donne des taux aux alentours des 70 % (en échographie). Les multicorrélations par ordre sont un peu moins performantes avec 65 % de moyenne où l'ordre quatre semble le plus discriminant. Sur ces résultats, l'égalisation d'histogramme semble améliorer les résultats (+ 3 % en moyenne). En combinant les ordres, les performances en mode supervisé des SOS atteignent le niveau de celles du SO. Par contre, nous remarquons que l'égalisation d'histogramme des échographies ne paraît plus avoir le même impact positif sur les résultats en combinant les ordres.

A présent, si nous nous concentrons sur les lignes du tableau 5.6.c pour comparer les paramètres de simulation (*cf.* figure 5.18), les lignes fournissant les meilleurs pourcentages correspondent aux simulations $S01/S09$ et $S05/S013$. La fusion des descripteurs (lignes indiquées par un « F »), qui a pour effet de stabiliser les descripteurs d’une glande (par moyennage), donne des résultats inférieurs aux autres simulations ce qui semble indiquer que le moyennage des descripteurs entraîne une perte trop importante d’information. L’uniformisation des descripteurs ou des images (lignes indiquées par un « N ») par acquisition ne semble pas non plus améliorer la classification des glandes mais dégrade les performances de façon moindre que la fusion.

Simulations / Descripteurs :				Images	SO	SO _{log}	c ₂	c ₃	c ₄	c _{1,2}	c _{2,4}	c _{1,4}	SO+ c ₂	SO+ c _{1,2}	SO+ c _{2,4}	SO+ c _{1,4}	SO _{log} + c ₂	SO _{log} + c _{1,2}	SO _{log} + c _{2,4}	SO _{log} + c _{1,4}
T	Img	Descrip.	N°																	
Échographie			S01	47	71	71	58	53	53	58	69	69	71	71	73	73	75	75	75	75
			F S02	38	73	76	67	62	65	67	72	74	73	73	74	74	77	76	77	77
		N	S03	47	69	69	56	55	55	59	66	66	70	69	69	68	72	71	69	69
		N	F S04	38	71	73	67	63	63	66	67	67	69	70	68	68	74	74	70	70
		N	S05	47	70	69	57	53	53	58	68	68	70	70	72	72	74	74	74	74
		N	F S06	38	72	75	67	62	65	66	72	72	73	72	74	74	75	76	77	76
		N	N S07	47	69	69	56	55	55	59	66	66	70	69	69	68	71	71	69	69
		N	N F S08	38	71	73	67	63	63	66	67	67	70	70	68	68	74	74	70	69

(a) 25 % des glandes consacrées à l’apprentissage

Simulations / Descripteurs :				Images	SO	SO _{log}	c ₂	c ₃	c ₄	c _{1,2}	c _{2,4}	c _{1,4}	SO+ c ₂	SO+ c _{1,2}	SO+ c _{2,4}	SO+ c _{1,4}	SO _{log} + c ₂	SO _{log} + c _{1,2}	SO _{log} + c _{2,4}	SO _{log} + c _{1,4}	
T	Img	Descrip.	N°																		
Échographie			S01	56	65	64	58	55	55	59	67	68	67	67	69	69	72	72	71	71	
			F S02	38	74	75	68	63	62	69	70	71	72	73	73	73	76	77	75	76	
		N	S03	56	62	62	57	56	57	57	65	66	64	64	66	66	66	66	63	63	
		N	F S04	38	65	70	65	63	59	65	64	64	66	66	65	65	69	69	67	67	
		N	S05	56	64	63	57	55	55	58	65	66	66	66	66	68	67	70	70	71	71
		N	F S06	38	71	72	68	63	62	69	70	70	71	72	72	72	75	75	75	75	
		N	N S07	56	62	63	57	56	57	57	65	66	64	64	66	66	66	66	66	63	63
		N	N F S08	38	65	70	65	63	59	66	65	64	66	66	65	65	69	69	67	67	

(b) 12.5 % des glandes consacrées à l’apprentissage

Tableau 5.7 – Taux de bonne classification par SMA en diminuant la proportion de glandes consacrées à l’apprentissage par rapport au tableau 5.6.c.

Le tableau 5.7 montre les conséquences d’une diminution de la proportion de glandes dans la base d’apprentissage à 25 % et 12.5 % sur les échographies (sans égalisation). Ces réductions de la base d’apprentissage entraînent une baisse de résultats de l’ordre de 5 % en divisant par deux sa taille, et de 10 % en divisant par quatre. Toutefois, nous pouvons remarquer que les taux de classification des lignes $S02$ et $S06$ affichent quasiment les mêmes valeurs quel que soit la proportion d’apprentissage (jusqu’à 77 % de bon classement). Cela semble indiquer que la fusion des descripteurs par acquisition permet d’accroître les performances de classification lorsque nous disposons de peu de données pour modéliser les classes, en offrant une meilleure stabilité des modèles, mais que les performances plafonnent en augmentant le nombre de données d’apprentissage (phénomène de sur-apprentissage). Dans la suite de cette sous-section, nous allons vérifier si nous observons les mêmes tendances, quant aux descripteurs et aux paramètres de simulation, sur les performances en classification non supervisée.

2. Classification par partitionnement :

Les tableaux 5.8.a et 5.8.b présentent les résultats de classification par k -moyennes et par DBSCAN. Le premier constat qui saute aux yeux est que la classification par DBSCAN ne fonctionne pas puisque les taux de classification sont constants autour des 60 % de bonne classification. Or, ce taux, correspond à la proportion de glandes normales dans la base et, en regardant dans le détail le classement par DBSCAN, nous pourrions nous apercevoir que l’ensemble des glandes est, en pratique, classé en *Normale*. Cela signifie que le paramétrage de cet algorithme, qui fonctionnait pour la base KTH-TIPS, n’est pas adapté à la base de glande et que des investigations supplémentaires sont donc nécessaires pour adapter cet algorithme aux caractéristiques de la base, ou pour faire évoluer cet algorithme de partitionnement par densité vers un algorithme plus performant comme OPTICS.

Avec l’algorithme des k -moyennes, nous obtenons des résultats plus satisfaisants qui sont proches de ceux de la classification par SMA. Nous pouvons remarquer que, comme en supervisé, le SO est plus performant pour décrire les échographies sans prétraitement alors que les multicorrélations

sont plus discriminantes après égalisation d’histogramme. Les ordres un et deux des descripteurs statistiques semblent être les ordres les plus discriminants avec les k -moyennes puisque nous atteignons 78 % de bonne classification. Contrairement au mode supervisé, le fait de combiner les deux types de descripteurs n’a pas forcément amélioré les résultats. Il serait toutefois intéressant de combiner les descripteurs du SO sans égalisation avec ceux des SOS avec égalisation. Comme pour le mode supervisé, nous remarquons que les simulations sans uniformisation ni fusion donnent globalement de meilleurs résultats même si, dans certains cas particuliers (SO, C_2 et $C_{1,2}$), la fusion améliore les performances. Enfin, nous pouvons noter une nette augmentation des taux de classification sur les élastographies puisque nous atteignons les 71 % sans uniformisation, ni fusion.

Simulations / Descripteurs :				Images	SO	SO _{log}	C ₂	C ₃	C ₄	C _{1,2}	C _{2,4}	C _{1,4}	SO + C ₂	SO + C _{1,2}	SO + C _{2,4}	SO + C _{1,4}	SO _{log} + C ₂	SO _{log} + C _{1,2}	SO _{log} + C _{2,4}	SO _{log} + C _{1,4}
T	Img	Descrip.	N°																	
Echographie			S01	48	75	67	64	62	42	64	59	59	75	75	64	66	69	69	67	67
			F S02	51	77	70	69	61	52	69	53	53	74	74	58	58	69	69	67	67
		N	S03	48	64	71	59	56	59	60	59	59	58	58	59	59	70	70	60	60
		N	F S04	51	69	69	61	52	53	61	57	57	60	60	57	57	66	64	57	57
		N	S05	47	76	69	61	62	42	61	56	56	58	58	57	62	68	68	68	68
		N	F S06	52	75	70	64	62	53	64	53	54	69	71	54	55	71	69	66	66
		N	N S07	47	64	71	59	56	59	60	59	60	58	58	58	60	70	70	60	60
		N	N	F S08	52	69	69	61	52	53	61	57	57	59	60	57	57	65	66	57
Echographie EH			S09	54	64	58	74	67	47	74	70	70	76	76	71	71	65	65	59	59
			F S10	58	61	55	78	54	53	78	54	55	71	71	59	59	60	60	57	59
		N	S11	55	50	50	69	52	46	69	47	47	61	61	47	47	62	59	47	47
		N	F S12	58	52	48	70	54	51	70	50	50	62	61	48	48	62	62	48	48
		N	S13	54	65	56	73	67	47	74	69	69	75	75	71	71	64	64	59	59
		N	F S14	60	60	51	75	54	53	78	54	54	73	72	58	58	56	56	55	56
		N	N S15	55	50	50	69	52	46	69	47	47	61	61	47	47	62	59	47	47
		N	N	F S16	59	52	48	70	54	51	70	50	50	62	61	48	48	62	62	48
Elastographie			S17	68	68	71	68	55	63	68	67	68	70	71	70	70	70	70	70	71
			F S18	64	70	68	70	47	60	70	66	59	68	68	67	68	68	69	69	
		N	S19	68	55	58	52	54	53	53	47	47	54	53	47	47	57	57	47	47
		N	F S20	64	47	50	50	51	53	50	47	47	47	47	47	47	51	52	47	47
		N	S21	51	58	51	53	55	63	54	52	52	51	53	51	51	53	53	52	52
		N	F S22	50	50	50	52	47	61	51	53	51	51	53	50	50	51	51	50	50
		N	N S23	51	55	58	52	53	53	53	47	47	54	53	47	47	57	57	47	47
		N	N	F S24	50	47	50	50	51	53	50	47	47	47	47	47	52	51	47	47

(a) Par k -moyennes (k -means)

Simulations / Descripteurs :				Images	SO	SO _{log}	C ₂	C ₃	C ₄	C _{1,2}	C _{2,4}	C _{1,4}	SO + C ₂	SO + C _{1,2}	SO + C _{2,4}	SO + C _{1,4}	SO _{log} + C ₂	SO _{log} + C _{1,2}	SO _{log} + C _{2,4}	SO _{log} + C _{1,4}	
T	Img	Descrip.	N°																		
Echographie			S01	62	61	59	61	62	62	62	62	62	62	62	61	61	62	62	62	62	
			F S02	61	63	62	61	62	62	62	61	61	62	62	62	62	62	62	62	62	
		N	S03	62	61	62	60	62	62	61	61	62	61	61	61	61	59	60	62	62	
		N	F S04	61	62	62	62	62	61	62	61	61	62	61	61	62	62	61	61		
		N	S05	62	62	59	61	62	62	61	62	62	62	62	62	62	62	62	62	62	
		N	F S06	61	64	62	61	62	62	62	61	62	63	63	62	62	62	62	62	62	
		N	N S07	62	61	62	60	62	62	61	61	62	61	61	61	61	59	60	62	62	
		N	N	F S08	61	62	62	62	62	61	62	61	61	61	62	61	61	62	62	61	61
Echographie EH			S09	62	61	62	62	62	62	62	61	62	62	62	61	61	62	62	62	62	
			F S10	62	62	61	62	62	62	62	62	62	62	62	62	62	62	62	62	62	
		N	S11	62	60	61	57	61	62	59	62	62	62	62	60	60	59	61	62	62	
		N	F S12	62	62	62	61	62	59	61	62	62	62	62	62	62	62	62	62	62	
		N	S13	62	61	60	61	62	62	62	62	62	62	62	61	61	61	62	62	62	
		N	F S14	62	62	61	62	62	62	62	62	62	62	62	62	62	60	60	62	62	
		N	N S15	62	60	61	57	61	62	59	62	62	62	62	60	60	59	61	62	62	
		N	N	F S16	62	62	62	61	62	59	61	62	62	62	62	62	62	62	62	62	62
Elastographie			S17	63	63	63	63	62	62	62	62	62	63	63	63	62	63	63	63	63	
			F S18	62	63	63	63	62	63	63	63	63	63	63	63	63	63	63	63	63	
		N	S19	63	63	63	62	62	63	63	62	62	63	63	62	62	63	63	62	62	
		N	F S20	62	63	63	63	62	60	63	62	62	62	63	63	62	63	63	63	63	
		N	S21	62	63	63	62	62	62	63	62	63	63	63	63	63	62	63	63	63	62
		N	F S22	63	62	63	63	62	63	63	63	63	63	63	63	63	63	63	63	63	63
		N	N S23	62	63	63	62	62	63	63	62	62	63	63	62	62	63	63	62	62	
		N	N	F S24	63	63	63	63	62	60	63	62	62	62	63	63	62	63	63	63	63

(b) Par DBSCAN

Tableau 5.8 – Taux de bonne classification pour deux algorithmes de partitionnement (*cf.* tableaux 5.7 pour plus d’informations sur les lignes et les colonnes).

3. Classification Spectrale (CS) :

En comparant les tableaux 5.8.a et 5.9, nous constatons que les performances de la classification spectrale (avec les k -moyennes) ne surpassent pas celles des k -moyennes, contrairement à ce que nous avons aperçu sur la base de référence. Le SO sans logarithme et les multicorrélations d’ordre deux (C_2 ou $C_{1,2}$) donne les meilleurs résultats avec respectivement 76 % (S02) et 78 % (S14). La combinaison des deux approches n’améliore pas les performances de classification, mais, comme pour les k -moyennes, l’association des descripteurs du SO sans égalisation d’histogramme des échographies avec ceux des SOS avec égalisation augmenterait sûrement les résultats de classification.

Simulations / Descripteurs :			Images	SO	SO _{log}	c ₂	c ₃	c ₄	c _{1,2}	c _{2,4}	c _{1,4}	SO + c ₂	SO + c _{1,2}	SO + c _{2,4}	SO + c _{1,4}	SO _{log} + c ₂	SO _{log} + c _{1,2}	SO _{log} + c _{2,4}	SO _{log} + c _{1,4}	
T	Img	Descrip.	N°																	
Echographie			S01	47	76	69	64	63	49	64	61	61	71	70	58	59	68	68	68	68
			F S02	50	75	71	68	60	48	68	55	55	73	73	50	50	69	69	66	67
		N	S03	47	64	71	61	53	62	61	57	57	61	61	57	57	69	71	57	57
		N	F S04	50	62	67	61	52	53	61	55	55	60	60	56	55	66	67	56	56
		N	S05	48	76	69	59	63	49	59	56	56	61	61	57	57	69	69	67	67
		N	F S06	52	76	69	61	60	48	61	56	56	64	64	50	50	68	68	66	66
		N	N S07	48	64	71	61	53	62	61	57	57	61	61	57	57	69	71	57	57
		N	N F S08	52	62	67	61	52	53	61	55	55	60	60	56	56	67	62	56	56
Echographie EH			S09	54	66	55	74	65	52	74	74	74	75	75	73	73	64	64	59	59
			F S10	52	60	51	77	57	55	77	55	55	70	70	59	59	57	57	58	58
		N	S11	54	50	50	68	53	49	68	50	50	59	59	50	50	49	49	49	49
		N	F S12	52	50	53	69	54	48	71	48	48	62	63	48	48	62	62	48	48
		N	S13	51	65	54	74	65	52	74	72	72	75	75	73	73	64	64	59	59
		N	F S14	52	61	51	78	57	55	78	54	54	69	69	57	57	54	53	57	57
		N	N S15	51	50	50	68	53	49	68	50	50	59	59	50	50	49	49	49	49
		N	N F S16	52	50	53	69	54	48	71	48	48	62	63	48	48	62	62	48	48
Elastographie			S17	64	68	67	60	53	61	60	58	59	62	63	63	62	68	68	68	68
			F S18	60	69	68	60	47	58	59	55	56	63	63	65	65	68	68	68	68
		N	S19	64	54	53	54	50	50	53	47	47	53	53	47	47	55	55	48	48
		N	F S20	60	47	48	51	51	50	52	48	48	50	50	47	47	50	50	47	47
		N	S21	49	54	51	50	53	62	49	55	53	50	50	53	52	53	53	52	52
		N	F S22	50	49	52	50	47	58	50	50	50	48	48	47	47	50	50	50	50
		N	N S23	49	54	53	54	50	50	53	47	47	53	53	47	47	55	55	48	47
		N	N F S24	50	47	48	51	51	50	52	48	48	50	50	47	47	50	50	47	47

Tableau 5.9 – Taux de bon classement des glandes par classification spectrale.

4. Focalisation sur les meilleures performances :

Dans ce paragraphe, nous allons revenir sur les simulations qui ont permis d’obtenir les meilleurs taux de bonne classification en projetant les descripteurs dans un plan à trois dimensions (méthode t-SNE) et en analysant leurs matrices de confusion, leurs sensibilités et spécificités.

La classification supervisée par SMA a atteint les 81 % de bonne classification en combinant les descripteurs SO_{log} et c₂ sans uniformisation par acquisition ni fusion, avec uniformisation globale par ordre et sur les échographies non égalisées (S01). La figure 5.19.d montre la projection dans un espace à trois dimensions de ces descripteurs mixtes. Nous observons que les glandes pathologiques (ronds noirs) sont plutôt situées sur la partie gauche du nuage de points et les glandes normales sur la partie droite. La frontière n’est néanmoins pas flagrante ce qui peut expliquer le taux plus faible des *k*-moyennes (68.5 %) par rapport à la SMA (80.6 %). Le tableau 5.10.d montre que la classification par SMA des descripteurs multirésolutions et statistiques permet d’obtenir une sensibilité de 81.2 % et une spécificité de 78.9 %, ce qui est proche des valeurs évoquées à la sous-section 5.2.1 : sensibilité moyenne de 77.9 % ou 71.3 %, et spécificité de 86.7 % ou 91.9 % (suivant l’article considéré).

En classification non supervisée, les descripteurs multirésolutions et statistiques se sont avérés plus discriminants en étant utilisés de manière séparée. Par *k*-moyennes et par CS, nous approchons les 78 % de bonne classification. Au travers de la figure 5.19 et du tableau 5.10, ce paragraphe propose de revenir en détail sur deux simulations particulières : (S02 - SO) et (S10 - c₂). La différence entre S02 et S10 réside dans le fait que les descripteurs ont été calculés sur les échographies, respectivement, sans et avec égalisation d’histogramme. Pour ces deux simulations, ni les images, ni les descripteurs n’ont été uniformisés par acquisition. En revanche, les descripteurs d’une même acquisition ont été fusionnés puis uniformisés par ordre par rapport à l’ensemble des descripteurs. En comparant la représentation 3D des images vectorisées (figure 5.19.a) par rapport à celles des deux simulations évoquées dans ce paragraphe (figures 5.19.b et 5.19.c), nous constatons que les glandes pathologiques et normales sont, comme le montre les taux de classification, devenues plus facilement séparables par leurs descripteurs multirésolutions ou statistiques. Nous retrouvons ce constat dans les matrices de confusion du tableau 5.10. Toutefois, nous pouvons constater que, en termes de sensibilité et de spécificité, la classification non supervisée semble moins performante que la SMA avec 50 % des glandes consacrées à l’apprentissage.

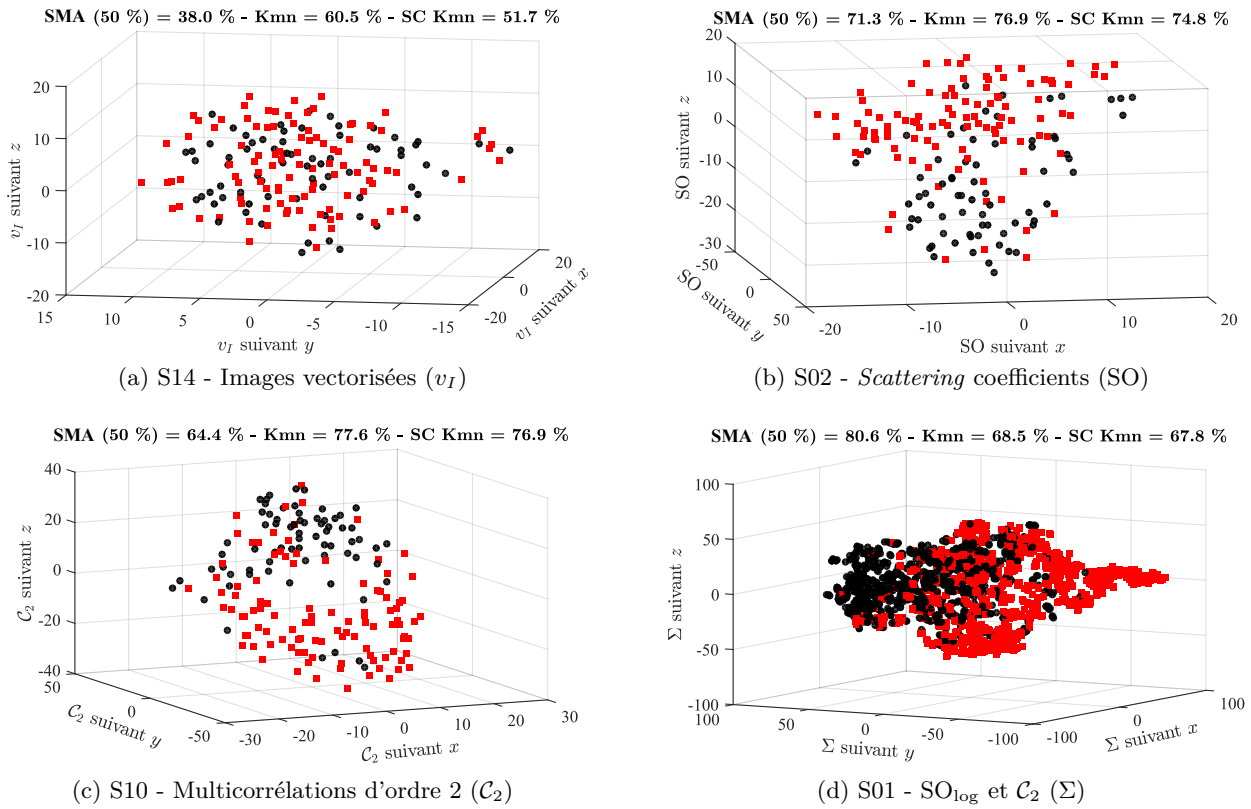


FIGURE 5.19 – *Manifold untangling* de la base des glandes salivaires par la méthode t-SNE pour quatre jeux de paramètres et de descripteurs : ces derniers sont projetés dans un espace à trois dimensions qui préservent les distances où chaque point représente un descripteur d’une imagerie (figure 5.19.d) ou d’une acquisition (car les descripteurs de chaque acquisition ont été fusionnés). Les points représentés par des ronds noirs (respectivement des carrés rouges) correspondent à la classe *Gougerot* (respectivement *Normale*). Les taux de bonne classification par SMA (50 % dédié à l’apprentissage), par k -moyennes (Kmn) et par CS ($SC Kmn$) sont indiqués au-dessus de chaque figure.

C_{kmn} (%)	Gougerot	Normale	C_{kmn} (%)	Gougerot	Normale	$C_{sc kmn}$ (%)	Gougerot	Normale	C_{sma} (%)	Gougerot	Normale
Gougerot	18.0	82.0	Gougerot	63.6	36.4	Gougerot	80.0	20.0	Gougerot	78.1	21.9
Normale	13.0	87.0	Normale	14.8	85.2	Normale	39.7	60.2	Normale	18.0	82.0
Sensibilité (%)		58.1	Sensibilité (%)		81.1	Sensibilité (%)		66.8	Sensibilité (%)		81.2
Spécificité (%)		48.5	Spécificité (%)		70.1	Spécificité (%)		74.8	Spécificité (%)		78.9

(a) S14 - Images vectorisées (v_I) par k -moyennes

(b) S02 - *Scattering* coefficients (SO) par k -moyennes

(c) S10 - Multicorrélations d'ordre 2 (C_2) par classification spectrale

(d) S01 - SO_{log} et C_2 (Σ) par SMA

Tableau 5.10 – Matrice de confusion (C), sensibilité et spécificité pour quatre jeux de paramètres, descripteurs et algorithme de classification.

5. Bilan sur la classification des glandes parotides et submandibulaires :

Au final, les simulations exposées ci-dessus montrent que les tendances observées sur la base de référence ne s’appliquent pas forcément sur la base de glandes salivaires. Tout d’abord, nous avons pu observer que le *scattering operator* ne surpasse plus les multicorrélations pour discriminer les glandes salivaires : la texture des glandes joue donc sûrement un rôle moindre pour leur classification que dans le cas des images synthétiques de matériaux. Ensuite, les résultats obtenus en supervisée, qui étaient excellents sur la base de référence, ne sont pas forcément meilleurs que les résultats obtenus

en non supervisé (en les regardant de manière globale). Cela indique que les classes de glandes sont probablement plus difficilement modélisables que celles de matériaux. Parmi les facteurs qui peuvent affecter négativement la classification, il y a :

- Le mélange des glandes parotides et submandibulaires (notamment pour les élastographies) ;
- Une frontière pas forcément évidente entre certaines imagettes extraites de glandes normales et pathologiques de grade un ;
- Des paramètres d'acquisition différents suivant les patients.

Toutefois, en se focalisant sur certaines simulations donnant des bons taux de classification, nous nous sommes aperçus que la classification supervisée donnait des résultats plus proches (sensibilité et spécificité) des études références en échographie (*cf.* section 5.2.1) que la classification non supervisée. Ces résultats sont donc prometteurs sachant qu'il est assez logique qu'ils soient inférieurs à ceux obtenus sur la base KTH-TIPS. Les résultats obtenus à partir des élastographies sont moins satisfaisants mais le fait de mélanger parotides et submandibulaires peut expliquer ces faibles taux de classification puisque les études références ont montré qu'il était compliqué de diagnostiquer une glande submandibulaire par élastométrie. Dans les prochaines sous-sections, nous allons considérer les deux types de glandes de manière indépendante et observer si cela se ressent au niveau de la classification.

5.2.3.2 Parotides

Dans cette sous-section, nous nous concentrons uniquement sur les glandes parotides : la base est constituée de 43 glandes normales et 29 glandes pathologiques. Les performances de classification sont données par le tableau 5.11.

En échographie, la classification supervisée (tableau 5.11) des descripteurs multirésolutions (SO) des parotides donne des résultats plus inégaux que sur les deux types de glandes mélangés (tableau 5.6.c) : environ 60 % au lieu de 70 % en moyenne sans fusion, environ 74 % au lieu de 72 % en moyenne avec fusion. Le taux maximal de bonne classification, qui était de 74 % avec les glandes mélangées, est de 75 % sur les parotides seules avec le SO. Au niveau des descripteurs, nous retrouvons les mêmes tendances avec une baisse des performances sans fusion mais une hausse importante avec fusion. Le taux maximal de 74 % (sans fusion) en mélangeant les glandes est passé à 77 % (avec fusion) en considérant seulement les parotides. Sur les deux types de descripteurs seuls, l'égalisation d'histogramme a permis d'améliorer légèrement les performances de classification alors qu'elle ne l'avait pas fait en mélangeant les glandes. En mélangeant les deux types de descripteurs, nous obtenons des pourcentages nettement moins bons sans fusion et légèrement supérieurs avec fusion. La valeur maximale de 81 % sans fusion sur les glandes mélangées n'est pas dépassée sur les parotides seules. Nous atteignons néanmoins les 78 % de bonne classification, ce qui, pour la simulation S_{10} avec les descripteurs $SO + \mathcal{C}_{2..4}$, correspond à une sensibilité de 82.3 % et une spécificité de 70.3 %.

Au niveau de la classification non supervisée des échographies, nous obtenons des résultats très intéressants avec des taux atteignant souvent les 80 % de bonne classification. L'algorithme des k -moyennes et la CS donne des résultats équivalents avec des pourcentages atteignant les 85 % de bon classement. Nous pouvons noter que les descripteurs $SO_{\log} + \mathcal{C}_2$ donnent des résultats stables (autour des 80 %) quels que soient les paramètres de simulations. La simulation S_{10} avec les descripteurs $SO + \mathcal{C}_{1..2}$ affiche 85 % de bonne classification avec les k -moyennes et la CS. Les matrices de confusions correspondantes montrent que cela représente une sensibilité de 84.4 % et une spécificité de 85.9 % (pour les deux algorithmes), ce qui est meilleur qu'en mélangeant les deux types de glandes.

En élastographie, la classification supervisée est en légère hausse (de 68 % au maximum à 72 %). En revanche, en classification non supervisée, nous obtenons des taux très intéressants allant jusqu'à 76 % de bonne classification avec les k -moyennes. En observant les projections 3D des élastographies de la figure 5.20, il est difficile d'établir une frontière claire entre les deux classes de points mais nous pouvons apercevoir que les parotides normales sont plutôt regroupées. Nous constatons également que, au niveau des projections 3D ou des taux de classification, l'apport des descripteurs multirésolutions et/ou statistiques est limité (2 à 3 %). Le meilleur taux obtenu par élastographie est 75.7 % (avec les k -moyennes) et affiche une sensibilité de 84.8% et une spécificité de 65.3 %, ce qui est proche des valeurs obtenues dans l'étude [83] (81 % et 67 %).

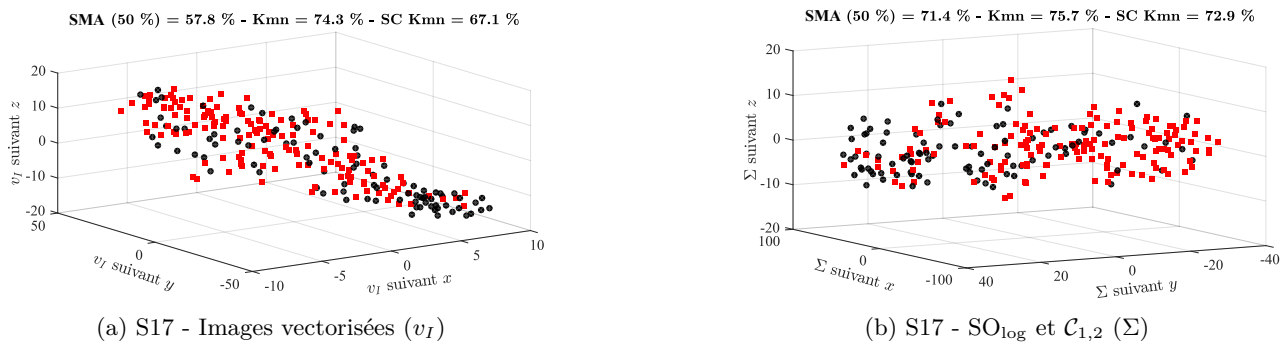


FIGURE 5.20 – *Manifold untangling* des élastographies des glandes parotides par la méthode t-SNE pour deux jeux de descripteurs (cf. figure 5.19).

5.2.3.3 Submandibulaires

A présent, nous allons nous focaliser uniquement sur les glandes submandibulaires : la base est également constituée de 44 glandes normales et 26 glandes pathologiques. Les performances de classification se trouvent dans le tableau 5.12.

Comme pour les échographies de parotides, nous observons des résultats plus inégaux en fonction des paramètres de simulation. Nous pouvons tout de même noter que, comme pour les glandes mélangées, les meilleurs taux sont obtenus pour les simulations S01 et S05, c'est-à-dire sans égalisation d'histogramme, ni fusion. De manière générale, la fusion des descripteurs par acquisition entraîne une baisse de résultats sur les submandibulaires, alors que c'est l'effet contraire pour les parotides. Cette différence peut provenir du fait que le nombre d'images extraites des glandes parotides, par leur taille plus importante, est beaucoup plus important que pour les submandibulaires (environ 1300 contre 700 images). Le taux maximal est de 82 % ce qui est légèrement supérieur aux taux maximaux des glandes mélangées (81 %) et des parotides seules (78 %). Ce taux maximal, acquis avec les multicorrélations d'ordre deux, trois et quatre ($C_{2..4}$) affiche une sensibilité bonne de 86.4 % et une spécificité plus faible de 74.0 % (pour S01).

En non supervisé, nous pouvons constater que les k -moyennes surpassent la CS avec des pourcentages atteignant les 86 % de bon classement (contre 83 % avec la CS). Les descripteurs statistiques, notamment l'ordre deux, semblent plus discriminants que les descripteurs multirésolutions. L'utilisation combinée des deux types descripteurs donne des résultats très intéressants (au-delà des 80 %) mais inférieurs aux 86 % de la corrélation croisée (C_2). Celle-ci permet d'établir une sensibilité de 87.7 % et une spécificité de 74.9 %. Nous remarquons également que l'uniformisation des descripteurs par acquisition dégrade la classification des glandes : ceci est aussi le cas, peut-être de manière moins flagrante, sur les parotides et le mélange des deux types de glande.

Enfin, au niveau des élastographies, nous retrouvons les constats des études présentées à la section 5.2.1 à savoir que l'élastographie ne semble pas permettre d'aider le diagnostic des glandes submandibulaires.

5.2.4 Bilan des simulations sur les glandes salivaires

Cette section a exposé une partie des simulations menées sur les glandes salivaires avec les approches d'extraction de descripteurs et de classification implémentées. Globalement, les résultats obtenus sont assez prometteurs puisque ces derniers sont dans le même ordre de grandeur que les études réalisées par la recherche médicale 5.2.1. En effet, nos simulations sur les échographies montrent que nous pouvons atteindre une sensibilité de 80 % (en supervisé ou en non supervisé), glandes mélangées. En revanche, nos résultats sont moins bons en terme de spécificité (quasiment 80 % en supervisé, 75 % en non supervisé et environ 90 % dans les études citées). En séparant les glandes parotides des submandibulaires, nous augmentons les performances des classifieurs non supervisés (k -moyennes et CS) qui montent jusqu'à 86 %. Sur les parotides, nous avons notamment atteint une sensibilité de 84.4 % et une spécificité de 85.9 %. Les résultats sur les submandibulaires affichent une meilleure sensibilité (87.7 %) mais une plus faible spécificité (74.9 %). En non supervisé, les résultats sont légèrement moins bons qu'en mélangeant les deux types de glandes mais cela vient peut-être du fait que, en les séparant, nous divisons la taille de la base de données par deux. Avec davantage de données, nous aurions peut-être eu une meilleure modélisation des classes, et donc une meilleure classification.

Ensuite, en ce qui concerne l'élastographie, nous constatons que, comme le montre la littérature, ce type d'imagerie donne des résultats plutôt satisfaisants sur les glandes parotides (mais en-dessous de l'échographie) mais que, actuellement, elle ne permet pas de discriminer les glandes submandibulaires pathologiques des normales. Les progrès en matière d'élastographie apportés par les nouveaux systèmes d'acquisition permettront, peut-être, à l'avenir de les dissocier. De plus, en termes d'approches, nous avons constaté que le *scattering operator* surpassait les statistiques d'ordre supérieur sur la base de référence (composée de textures). Sur la base de glandes, nous remarquons que les deux types de descripteurs donnent des résultats plus similaires, voire même meilleurs pour les descripteurs statistiques. Enfin, ces résultats ouvrent également de nombreuses perspectives d'amélioration comme la combinaison des échographies et des élastographies (pour les parotides), l'utilisation de nouveaux algorithmes de classification (comme l'apprentissage profond), une classification en un plus grand nombre de classes (grades d'avancée de la maladie), une estimation de la probabilité d'une glande à être pathologique (pourcentage par rapport aux imagerie étiquetées Gougerot), *etc.*

5.3 Classification du thrombus veineux

Après avoir validé nos approches sur une base de référence et obtenu des résultats prometteurs sur une base similaire, en termes de formation des images, à celle des thrombi, l'objectif des simulations présentées dans cette section est de trouver des corrélations entre le classement des images et le contexte de la Thrombose Veineuse Profonde (TVP) du patient. En se plaçant dans le cadre de la triade de Virchow (décrite à la section 1.1.2), nous pourrions espérer trouver trois classes correspondant à chacun des types de phénomènes engendrant la TVP :

- la **stase sanguine** due à une immobilisation prolongée (long voyage, plâtre, hospitalisation) ou à une mobilité réduite (due à l'âge par exemple) ;
- l'**altération des parois** veineuses à cause d'une intervention chirurgicale, de la pose d'un cathéter ou tout simplement du vieillissement des veines ;
- la **thrombophilie** c'est-à-dire un état d'hypercoagulabilité du sang dû à des prédispositions génétiques, des hormones ou un cancer.

A travers ces trois types de facteurs, nous pouvons comprendre toute la difficulté qui existe pour définir nos classes. Un cancer peut par exemple engendrer un ralentissement du flux sanguin et une hypercoagulabilité. De plus, certains patients présentent des TVP idiopathiques ce qui signifie que le principal facteur déclencheur n'a pas pu être identifié à partir de l'historique du patient. Toutefois, cela ne signifie pas que ce déclencheur n'existe pas. Potentiellement, ces patients pourraient donc souffrir d'un cancer non déclaré qui aurait favorisé l'apparition d'un thrombus [57]. Les TVPs dues à des immobilisations devraient se distinguer car les thrombi sont souvent plus anciens (une quinzaine de jours) et donc plus échogènes. Enfin, un partitionnement en fonction de la présence d'EP serait également un résultat positif. Néanmoins, la complexité réside dans le fait que nous ne pouvons pas affirmer qu'une TVP diagnostiquée sans EP n'aurait pas pu évoluer (sans le traitement anticoagulant), à terme, vers un épisode thromboembolique.

Cette section s'articule en six sous-sections. Elle commence par effectuer un rappel sur la base et le protocole expérimental utilisé pour obtenir les résultats présentés dans les cinq sous-sections suivantes. Les résultats sont divisés en fonction de l'objectif de classification :

1. Identifier le principal facteur déclencheur ;
2. Identifier la présence d'un facteur déclencheur et d'une embolie pulmonaire ;
3. Séparer les thromboses provoquées (avec un facteur déclencheur) des thromboses idiopathiques ;
4. Détecter la présence d'embolie pulmonaire.

Ensuite, la sixième et dernière sous-section fait le bilan des simulations menées sur les thrombi veineux.

5.3.1 Base d'étude et protocole expérimental

Les simulations présentées tout au long de cette section 5.3 ont été réalisées sur la base de thrombi veineux collectée avec le protocole 3 (*cf.* sous-section 1.3.2.3) pour des raisons qualitatives et quantitatives. Par ce protocole, les échographies et les élastographies sont acquises avec la version la plus récente du système Toshiba Aplio 500 (post-mise à jour) et avec la sonde convexe. La mise à jour a permis d'améliorer la qualité des élastographies par rapport au protocole 1. La sonde convexe a permis d'accroître leur stabilité (moins de variations entre les élastographies d'un même patient capturées par un même praticien) par rapport à la sonde linéaire utilisée aux protocoles 2A et 2B. De plus, c'est avec ce protocole 3 que nous avons collecté le plus grand nombre d'images. Toutefois, les résultats préliminaires sur les bases obtenues avec les protocoles antérieurs sont disponibles dans nos publications ([9, 10, 11] pour le protocole 1, [12, 13] pour le protocole 2A) et en annexe F.

La base collectée avec le protocole 3, et étudiée dans cette section, est résumée en figure 5.21 : les identifiants (ID) des patients sont regroupés en fonction du contexte d'apparition de la TVP et

de la présence d'une EP. L'idéal serait, à travers les simulations présentées par la suite, de retrouver les facteurs déclencheurs ou les quatre classes de la figure 5.21 qui dépendent de deux critères : le caractère provoqué (facteur déclencheur connu) de la TVP et la présence d'EP. A travers les quelques images affichées en figure 5.21 (une par mode d'imagerie et pour chacune des quatre classes), nous pouvons constater qu'il paraît difficile, à l'œil nu, de caractériser la structure des thrombi. Au moyen du *scattering operator* et des statistiques d'ordre supérieurs, nous espérons extraire des descripteurs de ces images qui pourront nous permettre de discriminer les images en fonction de l'épidémiologie (facteurs de risque, EP) de la thrombose.

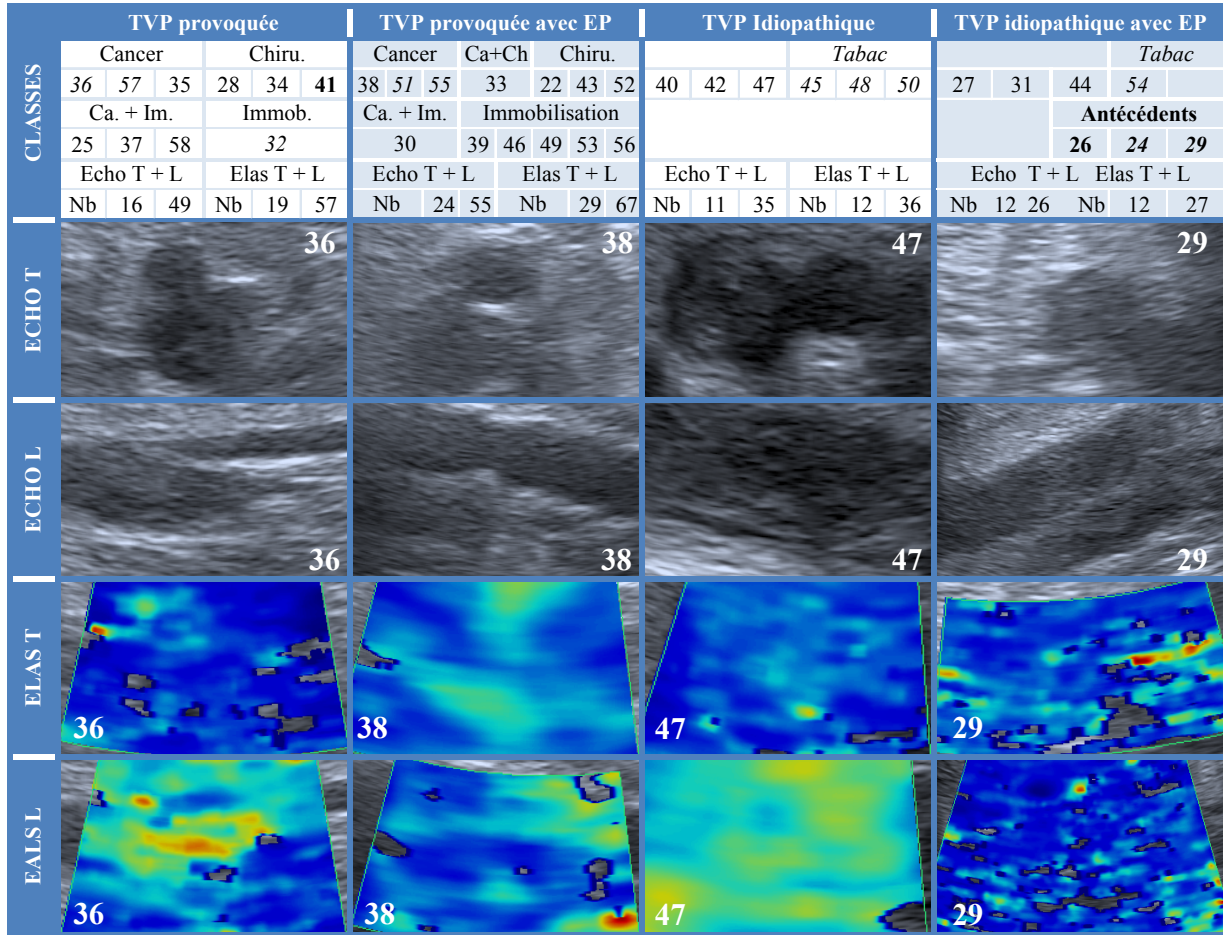


FIGURE 5.21 – Base de thrombi organisée en quatre classes suivant la connaissance (TVP provoquée) ou non (TVP idiopathique) du facteur déclencheur et la présence ou non EP : les IDs de chaque patient sont triés en fonction du facteur déclencheur ; les **IDs** en gras indiquent que le patient a déjà eu des antécédents de TVP ; les *IDs* en italique signifient que le patient fume ; pour chaque classe, il est indiqué le nombre d'images collectées par type (échographie et élastographie) et par coupe (transversale et longitudinale).

Le protocole expérimental appliqué pour classer les images de thrombi est identique à celui utilisé pour la base de glandes salivaires (résumé en sous-section et en figure 5.18). Nous noterons tout de même quelques différences. Les glandes pouvaient être séparées en fonction de leur type (parotide ou submandibulaire), ce qui n'est pas le cas des thrombi. Toutefois, nous disposons à présent de deux types d'images, transversal et longitudinal. Les résultats présentés dans les sections suivantes ont été obtenus en mélangeant ces deux types d'images : un thrombus est donc représenté par ses acquisitions transversales et longitudinales. Comme pour les glandes salivaires, un certain nombre d'images sont extraites de ces acquisitions et étiquetées en trois catégories « Thrombus » (au lieu de « Glande »), « Mixte » ou « Réf. ». Les images, en échographie et en élastographie, possèdent une taille de 32×32 pixels (taille la plus adaptée par rapport à la taille des thrombi). Les descripteurs

multirésolutions (quatre échelles, huit orientations, deux ordres de décomposition) et statistiques (de l'ordre un à quatre suivant les axes diagonaux, vertical et horizontal) sont calculés suivant les mêmes paramètres que pour les glandes salivaires. Les imagerie extraites, ou leurs descripteurs, peuvent être uniformisées ou fusionnées par acquisition (cela dépend du numéro de la simulation considérée). Enfin, le classement des descripteurs est également réalisé dans les mêmes conditions que pour les glandes salivaires : par Sélection de Modèles Affines (SMA) avec 50 % des thrombi consacrés à l'apprentissage (règle de décision \mathcal{R}_2 , cf. section 5.2.2.5), par k -moyennes et par classification spectrale⁴.

5.3.2 Facteurs déclencheurs des TVP

Dans cette sous-section, nous souhaitons classer les thrombi de notre base de données en fonction de leurs facteurs déclencheurs. A partir de la figure 5.21, nous distinguons donc cinq classes :

1. *Chirurgie* : patients 22, 28, 34, 41, 43 et 52 ;
2. *Génétique* (soit, pour cette base, avec antécédents familiaux) : patients 24, 26 et 29 ;
3. *Cancer* : patients 25, 30, 33, 35, 36, 37, 38, 51, 55, 57 et 58 ;
4. *Idiopathique* : patients 27, 31, 40, 42, 44, 45, 47, 48, 50 et 54 ;
5. *Immobilisation* : patients 32, 39, 46, 49, 53 et 56.

Nous pouvons déjà noter que certains patients pourraient appartenir à deux classes différentes puisque, par exemple, quatre des patients de la base ont déclaré une TVP suite à une immobilisation et en présence d'un cancer. Pour les simulations présentées dans cette sous-section, ces patients sont attribués à la classe *Cancer* car nous pensons que le cancer a peut-être plus d'influence sur la structure du thrombus que l'immobilisation (ou la chirurgie pour le patient 33).

Les résultats de classification suivant les cinq classes évoquées ci-dessus sont donnés dans les tableaux 5.16. Le tableau 5.16.a affiche les taux de bonne classification en consacrant 50 % des thrombi de chaque classe à l'apprentissage (pour 10 partitions aléatoires de chaque classe). Par exemple, la classe *Chirurgie* est modélisée à partir des données de trois patients et les données restantes sont utilisées pour le test. Les lignes du tableau 5.16.a correspondent aux paramétrages de la simulation (type d'image, uniformisation des imagerie par acquisition, uniformisation et fusion des descripteurs des imagerie par acquisition) et les colonnes indiquent les vecteurs descripteurs utilisés (images vectorisées, *scattering* coefficients SO, multicorrélations \mathcal{C} et descripteurs combinés).

C_{SMA}	Ch.	Gé.	Ca.	Id.	Im.
Ch.	0.0	0.0	33.3	66.7	0.0
Gé.	0.0	0.0	80.0	20.0	0.0
Ca.	0.0	0.0	58.0	42.0	0.0
Id.	1.7	1.7	21.7	75.0	0.0
Im.	0.0	0.0	16.7	76.7	6.7

(a) Elastographie - S18 - $SO_{log} + \mathcal{C}_2$ (42.2 %)

C_{SMA}	Ch.	Gé.	Ca.	Id.	Im.
Ch.	0.0	0.0	26.7	73.3	0.0
Gé.	0.0	0.0	20.0	80.0	0.0
Ca.	0.0	0.0	42.0	58.0	0.0
Id.	0.0	0.0	25.0	75.0	0.0
Im.	0.0	0.0	50.0	50.0	0.0

(b) Echographie - S02 - $SO_{log} + \mathcal{C}_2$ (38.8 %)

Tableau 5.13 – Matrice de confusion (pourcentage moyen sur 10 partitions) par SMA en fonction des facteurs déclencheurs des thrombi veineux : chirurgie (Ch.), génétique (Gé.), cancer (Ca.), idiopathique (Id.) et Immobilisation (Im.).

En observant le tableau 5.16.a, nous constatons que chaque simulation (quel que soit le paramétrage et la définition des descripteurs) donne des taux relativement faibles (32 % en moyenne) qui oscillent entre 26 et 42 %. Le maximum est obtenu pour la simulation S18 (élastographie avec fusion des descripteurs par acquisition) avec les descripteurs combinés ($SO_{log} + \mathcal{C}_2$). La matrice de confusion

4. Comme pour les glandes salivaires, le DBSCAN n'a pas permis d'aboutir à des résultats concluants : soucis de paramétrage et/ou algorithme non adapté à ce type de base.

correspondant à ce taux maximal (tableau 5.13.a) montre que la majorité des élastographies sont attribuées aux classes *Cancer* et *Idiopathique*. Ces résultats pourraient indiquer que les thrombi d'origine cancéreuse et génétique auraient une structure différente des thrombi idiopathiques, dus à une immobilisation ou à la chirurgie. Toutefois, la matrice de confusion de la même simulation appliquée sur les échographies (tableau 5.13.b) affiche une répartition différente des thrombi dans ces deux mêmes classes.

C_{kmn}	Ch.	Gé.	Ca.	Id.	Im.
Ch.	3.3	1.7	61.7	23.3	10.0
Gé.	6.7	60.0	30.0	0.0	3.3
Ca.	2.0	2.0	49.0	32.0	15.0
Id.	1.5	23.1	13.8	42.3	19.2
Im.	0.0	0.0	15.7	24.3	60.0

 (a) Elastographie - $S22$ - $SO_{\log} + C_2$ (42.6 %)

C_{kmn}	Ch.	Gé.	Ca.	Id.	Im.
Ch.	58.3	0.0	13.3	20.0	8.3
Gé.	0.0	33.3	50.0	16.7	0.0
Ca.	29.0	0.0	48.0	3.0	20.0
Id.	26.9	6.2	19.2	39.2	8.5
Im.	32.0	10.0	14.0	20.0	24.0

 (b) Echographie - $S01$ - $SO_{\log} + C_{1,2}$ (42.2 %)

Tableau 5.14 – Matrice de confusion (pourcentage moyen sur 10 tests) par k -moyennes en fonction des facteurs déclencheurs des thrombi veineux.

Dans le paragraphe précédent consacré aux résultats de la classification supervisée, nous avons constaté qu'une classification en deux classes serait certainement plus adaptée (qu'en cinq classes) pour notre base de thrombi. En s'intéressant au tableau 5.16.b, nous allons vérifier cette première tendance sur les résultats des k -moyennes (toujours en cinq classes). Contrairement à ce que pourrait laisser penser l'échelle des couleurs, les résultats des k -moyennes sont similaires à ceux de la SMA (pourcentage moyen de 33 %) puisqu'ils oscillent entre 24 et 43 %. En se focalisant sur les matrices de confusion de deux des meilleurs taux de classification (tableaux 5.14), nous observons des résultats totalement différents de ceux de la classification supervisée. Dans les deux cas, nous observons une confusion importante entre les classes *Chirurgie* et *Cancer*. En élastographie (Simulation $S22$), les thrombi dus à une immobilisation sont plutôt bien classés (60 %), alors qu'en échographie ($S01$), ils sont répartis parmi l'ensemble des *clusters*.

C_{cs}	Ch.	Gé.	Ca.	Id.	Im.
Ch.	28.3	11.7	41.7	16.7	1.7
Gé.	16.7	56.7	26.7	0.0	0.0
Ca.	29.0	8.0	39.0	11.0	13.0
Id.	9.2	7.7	2.3	65.4	15.4
Im.	0.0	0.0	10.0	50.0	40.0

 (a) Elastographie - $S17$ - $SO_{\log} + C_2$ (47.7 %)

C_{cs}	Ch.	Gé.	Ca.	Id.	Im.
Ch.	33.3	6.7	13.3	26.7	20.0
Gé.	0.0	33.3	60.0	0.0	6.7
Ca.	20.0	10.0	40.0	20.0	10.0
Id.	36.9	9.2	0.0	46.2	7.7
Im.	20.0	20.0	20.0	0.0	40.0

 (b) Echographie - $S07$ - C_3 (40.5 %)

Tableau 5.15 – Matrice de confusion par classification spectrale en fonction des facteurs déclencheurs des thrombi veineux.

En revanche, le tableau 5.16.c semble indiquer que la classification spectrale est légèrement plus performante, pour classer les élastographies de thrombi, que les deux algorithmes évoqués précédemment. La combinaison des descripteurs multirésolutions et statistiques permet d'atteindre 48 % de bonne classification. La matrice de confusion correspondante, affichée dans le tableau 5.15.a, montre que quatre des cinq classes sont moyennement classées : la majorité des TVP idiopathiques sont groupées ensemble, de même que pour les TVP avec antécédents génétiques. En échographies, les taux de classification sont un peu inférieurs : le taux maximal est de 40 %. La matrice de confusion correspondante est donnée par le tableau 5.15.b et montre que chaque classe est bien classée avec au mieux 46 % (sur la diagonale de la matrice). Nous pouvons néanmoins noter que 60 % des TVP associées à un cancer sont regroupées avec les TVP avec antécédents génétiques.

Finalement, les résultats présentés dans cette section semblent indiquer que la classification en cinq classes, définies par rapport aux facteurs déclencheurs, paraît compliquée. La classification supervisée par SMA suggère plutôt une classification en deux classes et la classification non supervisée en trois ou quatre classes (il y a au moins un *cluster* qui ne comprend pas beaucoup de points). Les taux de classification, relativement bas (environ 40 %), peuvent signifier que :

- Le facteur déclencheur n’a pas un impact visible sur la structure du thrombus ;
- Cet impact du facteur déclencheur n’est pas visible par échographie et/ou élastographie ;
- Les approches de descriptions des images et/ou les classifieurs utilisés ne sont pas adaptés aux données de thrombi : les frontières entre les classes sont peut-être non linéaires, ce qui expliquerait la difficulté de modélisation des classes par SMA ;
- Les variations des paramètres d’acquisition (comme le gain) altèrent les résultats de classification ;
- L’âge des thrombi (pas forcément connu) n’est pas pris en compte pour la classification : deux thrombi provoqués par le même facteur déclencheur mais à des âges différents (une semaine d’intervalle par exemple) auront probablement des structures différentes ;
- Les cinq classes choisies ne sont pas suffisamment bien définies (nous avons notamment vu que certains thrombi pourraient en effet être classés dans différentes classes).

Toutefois, quel que soit la (ou les) raison(s) de ces taux assez faibles, ces résultats sont à relativiser dans le sens où la base actuelle est encore de taille limitée (une trentaine de patients), notamment pour bien modéliser les classes en mode supervisé. Dans les prochaines sections, nous allons redéfinir nos classes et les simplifier afin d’observer si cela nous permet d’améliorer les performances de classification.

5.3.3 Caractère provoqué de la TVP et présence d’EP

Dans cette section, nous cherchons à classer les données de notre base de thrombi en quatre classes. Nous distinguons ainsi les TVP suivant deux critères : l’identification du facteur déclencheur principal (TVP provoquée/idiopathique) et la complication de la TVP en EP. Les résultats de classification, en fonction des paramètres de simulations et du choix des descripteurs, sont donnés dans le tableau 5.18.

En supervisé, les taux de bonne classification varient entre 25 et 45 % pour un taux moyen de 35 % (*cf.* tableau 5.18.a). La modélisation des quatre classes souhaitées paraît à nouveau compliquée par SMA. Nous pouvons noter que les taux obtenus à partir des élastographies sont une nouvelle fois supérieurs à ceux obtenus à partir des échographies. De plus, en analysant les matrices de confusion des deux simulations qui ont permis d’atteindre les taux maximaux en élastographie (45 % pour $S22 - SO_{\log} + C_{1,2}$) et en échographie (41.7 % pour $S06 - C_4$), nous observons que ces taux sont trompeurs et que, en réalité, la grande majorité des thrombi sont classés en *Idiopathique avec EP*. La SMA ne parvient donc pas à modéliser correctement les quatre classes souhaitées.

C_{SMA}	Prov.	Pr. EP	Idiop.	Id. EP
Prov.	0.0	0.0	2.5	97.5
Pr. EP	0.0	0.0	26.7	73.3
Idiop.	0.0	0.0	30.0	70.0
Id. EP	0.0	0.0	1.4	98.6

(a) Elastographie - $S20 - SO_{\log} + C_{1,2}$ (45.0 %)

C_{SMA}	Prov.	Pr. EP	Idiop.	Id. EP
Prov.	30.0	0.0	7.5	62.5
Pr. EP	3.3	0.0	10.0	86.7
Idiop.	0.0	10.0	15.0	75.0
Id. EP	3.3	0.0	8.3	88.3

(b) Echographie - $S06 - C_4$ (41.7 %)

Tableau 5.17 – Matrice de confusion par SMA en fonction du caractère provoqué de la TVP et de la présence d’EP : provoquée (Prov.), provoquée avec EP (Pr. EP), idiopathique (Idiop.) et idiopathique avec EP (Id. EP).

En non supervisé, l'algorithme des k -moyennes affiche (*cf.* tableau 5.18.b) des performances similaires (35 % en moyenne avec des taux compris entre 25 % et 46 %). Contrairement aux résultats précédents, les simulations menées sur les échographies donnent de meilleurs taux de bonne classification que celles réalisées sur les élastographies. En étudiant deux simulations en particulier ($S17 - SO_{\log} + C_2$ et $S07 - SO_{\log} + C_2$), leurs matrices de confusion, illustrées dans le tableau 5.19, affichent des résultats plus intéressants que celles présentées pour la SMA. La simulation $S07$ (sur les échographies) a permis de regrouper 73.4 % des thromboses provoquées ensemble ($73.4 = (57.8 + 55.7 + 13.3 + 20.0)/2$) et 57.25 % des thromboses idiopathiques ensemble. Il y a par contre une confusion importante entre les classes avec ou sans EP. En revanche, au niveau des élastographies, la simulation $S17$ affiche un bon pourcentage de classification des thromboses idiopathique (sans EP) mais les trois autres classes sont plus mélangées.

C_{kmn}	Prov.	Pr. EP	Idiop.	Id. EP
Prov.	8.9	0.0	10.0	81.1
Pr. EP	2.9	41.4	15.7	40.0
Idiop.	2.5	0.0	70.0	27.5
Id. EP	5.3	0.0	43.3	51.3

(a) Elastographie - $S17 - SO_{\log} + C_2$ (43.6 %)

C_{kmn}	Prov.	Pr. EP	Idiop.	Id. EP
Prov.	57.8	13.3	0.0	28.9
Pr. EP	55.7	20.0	14.3	10.0
Idiop.	23.8	32.5	17.5	26.3
Id. EP	18.5	10.8	1.5	69.2

(b) Echographie - $S07 - SO_{\log} + C_2$ (45.9 %)

Tableau 5.19 – Matrice de confusion par k -moyennes en fonction du caractère provoqué de la TVP et de la présence d'EP.

Ensuite, en ce qui concerne la classification spectrale dont les résultats sont illustrés dans le tableau 5.18.c, les performances obtenues paraissent à nouveau supérieures à celles des deux autres méthodes de classification utilisées. Elle permet d'obtenir jusqu'à 48 % de classification correcte sur les élastographies et jusqu'à 49 % sur les échographies. Les deux simulations détaillées par le tableau 5.20 montrent que plus de 60 % des thromboses provoquées et idiopathiques sont correctement séparées. L'identification de la présence d'EP est à nouveau source de confusion pour ces deux simulations.

C_{cs}	Prov.	Pr. EP	Idiop.	Id. EP
Prov.	66.7	15.6	17.8	0.0
Pr. EP	20.0	37.1	14.3	28.6
Idiop.	26.3	7.5	40.0	26.3
Id. EP	10.7	20.0	24.0	45.3

(a) Elastographie - $S23 - SO_{\log} + C_{2..4}$ (47.7 %)

C_{cs}	Prov.	Pr. EP	Idiop.	Id. EP
Prov.	61.1	27.8	0.0	11.1
Pr. EP	50.0	35.7	14.3	0.0
Idiop.	25.0	25.0	25.0	25.0
Id. EP	7.7	23.1	7.7	61.5

(b) Echographie - $S07 - SO_{\log} + C_2$ (48.7 %)

Tableau 5.20 – Matrice de confusion par classification spectrale en fonction du caractère provoqué de la TVP et de la présence d'EP.

Globalement, les quatre classes définies pour les simulations présentées dans cette section semblent encore relativement compliquées à dissocier quel que soit le type de descripteurs et l'algorithme de classification choisi. Toutefois, l'analyse des matrices de confusion des meilleures simulations (par classifieur et type d'image) est encourageante puisque nous avons pu constater que les TVP provoquées et idiopathiques étaient assez bien séparées et que la confusion se trouvait plutôt pour la détection de la présence d'EP. Pour étayer ce constat, la prochaine section considérera uniquement deux classes : provoquée ou idiopathique.

5.3.4 Caractère provoqué de la TVP

Dans cette section, l'objectif est de dissocier les TVP provoquées des TVP idiopathiques. Autrement dit, nous cherchons à identifier si le fait de posséder un facteur déclencheur identifiable (chirurgie, immobilisation, cancer, *etc.*) confère au thrombus une structure différente que s'il a été formé sans raison apparente. Une partie des résultats de classification (avec ces deux classes) est donnée au tableau 5.24. En observant le tableau 5.24.a, nous constatons que la SMA ne parvient pas à modéliser correctement les deux classes souhaitées puisque les taux de bonne classification varient entre 44 % et 60 %. Les matrices de confusion, correspondant aux deux meilleures simulations par élastographie et échographie, montrent que la majorité des thrombi est étiquetée idiopathique.

C_{SMA}	Provoquée	Idiopathique
Provoquée	33.8	66.3
Idiopathique	20.9	79.1

(a) Elastographie - $S17 - C_2$ (60.0 %)

C_{SMA}	Provoquée	Idiopathique
Provoquée	17.5	82.5
Idiopathique	7.0	93.0

(b) Echographie - $S01 - SO + C_{1,2}$ (59.4 %)

Tableau 5.21 – Matrice de confusion par SMA entre TVP provoquée et idiopathique.

En revanche, les deux techniques de classification non supervisées (les k -moyennes au tableau 5.24.b et la classification spectrale au tableau 5.24.c) fournissent de bien meilleurs résultats, notamment en échographie. En effet, en se focalisant sur deux simulations (*cf.* tableaux 5.22.a et 5.23.a), nous remarquons que la majorité des thrombi sont regroupés dans un même *cluster* (d'où environ 65 % de bonne classification). En revanche, en échographie, les matrices de confusion des deux meilleures simulations de chaque algorithme (tableaux 5.22.b et 5.23.b) montrent qu'environ 70 % des thrombi sont bien partitionnés, ce qui est plutôt satisfaisant.

C_{knn}	Provoquée	Idiopathique
Provoquée	20.6	79.4
Idiopathique	0.0	100.0

(a) Elastographie - $S17 - SO + C_2$ (67.4 %)

C_{knn}	Provoquée	Idiopathique
Provoquée	63.8	36.3
Idiopathique	25.7	74.3

(b) Echographie - $S06 - SO_{\log}$ (69.7 %)

Tableau 5.22 – Matrice de confusion par k -moyennes entre TVP provoquée et idiopathique.

C_{cs}	Provoquée	Idiopathique
Provoquée	46.3	53.8
Idiopathique	20.9	79.1

(a) Elastographie - $S21 - C_2$ (65.6 %)

C_{cs}	Provoquée	Idiopathique
Provoquée	75.0	25.0
Idiopathique	31.0	69.0

(b) Echographie - $S06 - SO_{\log} + C_{1..4}$ (71.6 %)

Tableau 5.23 – Matrice de confusion par classification spectrale entre TVP provoquée et idiopathique.

En analysant plus précisément les tableaux 5.24.b et 5.24.c, nous constatons que les meilleurs taux sont obtenus avec les descripteurs multirésolutions (avec le logarithme pour les k -moyennes, sans logarithme pour la CS), les multicorrélations d'ordre deux (C_2 ou $C_{1,2}$ pour la CS) et/ou avec les descripteurs combinés. Les multicorrélations d'ordres trois et quatre ne semblent pas pertinentes pour discriminer les thromboses provoquées des thromboses idiopathiques par échographie. Au niveau des paramètres de simulation, l'uniformisation des descripteurs multirésolutions par acquisition semble améliorer les performances de classification non supervisée. Les descripteurs statistiques paraissent mieux optimisés s'ils sont fusionnés et si les imagerie sont, en amont, uniformisées par acquisition.

5.3.5 Présence d'EP

Dans la section précédente, nous sommes parvenus à dissocier les thromboses provoquées des thromboses idiopathiques par échographies avec une réussite d'environ 70 %. A présent, nous allons nous intéresser uniquement à la présence, ou l'absence, d'Embolie Pulmonaire (EP) associée à la TVP. Les résultats de classification sont donnés par le tableau 5.28.

C_{SMA}	Sans EP	Avec EP
Sans EP	67.5	32.5
Avec EP	34.5	65.5

(a) Elastographie - $S24$ - $SO + C_{2..4}$ (66.3 %)

C_{SMA}	Sans EP	Avec EP
Sans EP	63.8	36.3
Avec EP	36.0	64.0

(b) Echographie - $S05$ - $C_{1..4}$ (63.8 %)

Tableau 5.25 – Matrice de confusion par SMA en fonction de la présence d'EP.

Le tableau 5.28.a montre que la classification supervisée par SMA donne des résultats plus satisfaisants que lors des simulations présentées aux sections précédentes. Les descripteurs statistiques et/ou combinées permettent de détecter la présence d'EP avec une précision d'environ 65 % (échographie et élastographie). Les matrices de confusion des simulations $S24$ et $S05$ (*cf.* tableau 5.25) confirment ces résultats puisque leurs diagonales affichent des taux proches des 65 %.

C_{knn}	Sans EP	Avec EP
Sans EP	76.5	23.5
Avec EP	48.2	51.8

(a) Elastographie - $S24$ - SO_{log} (62.6 %)

C_{knn}	Sans EP	Avec EP
Sans EP	54.7	45.3
Avec EP	10.0	90.0

(b) Echographie - $S07$ - SO (73.8 %)

Tableau 5.26 – Matrice de confusion par k -moyennes en fonction de la présence d'EP.

C_{cs}	Sans EP	Avec EP
Sans EP	71.8	28.2
Avec EP	43.6	56.4

(a) Elastographie - $S23$ - $SO_{log} + C_2$ (63.1 %)

C_{cs}	Sans EP	Avec EP
Sans EP	65.9	34.1
Avec EP	15.0	85.0

(b) Echographie - $S14$ - $C_{1,2}$ (76.2 %)

Tableau 5.27 – Matrice de confusion par classification spectrale en fonction de la présence d'EP.

En non supervisé (les k -moyennes au tableau 5.28.b et la classification spectrale au tableau 5.28.c), l'échogénicité du thrombus semble plus à même de permettre la détection de la présence d'EP que sa carte de dureté (jusqu'à 76 % de bon classement pour les échographies et jusqu'à 63 % pour les élastographies). Les matrices de confusion des tableaux 5.26 et 5.27 indiquent que les élastographies permettent de mieux regrouper les thrombi non associés à des EP (plus de 70 %) et que les échographies sont plus efficaces pour détecter la présence d'EP (plus de 85 %). La classification spectrale des multicorrélations d'ordre deux ($S14 - C_{1,2}$) affiche des performances prometteuses avec 76 % de bon classement. Sur ces simulations, le SO affiche des meilleurs résultats lorsqu'il est associé aux k -moyennes. En revanche, ces simulations ainsi que celles de la section 5.3.4 suggèrent que les multicorrélations (d'ordre deux notamment) sont mieux exploitées par la classification spectrale que par les k -moyennes.

5.3.6 Bilan des simulations sur les thrombi veineux

Dans cette section, nous nous sommes intéressés à la classification des images de thrombi en suivant le même protocole expérimental utilisé sur la base des glandes salivaires : même définition des descripteurs, mêmes paramètres de simulation (uniformisation des imagerie et des descripteurs par acquisition, fusion des descripteurs par acquisition) et mêmes algorithmes de classification. Tout d'abord, nous avons cherché à identifier les facteurs déclencheurs des thromboses veineuses profondes (chirurgie, facteur génétique, cancer, immobilisation et idiopathique) par échographie et élastographie. Nous avons constaté que les résultats obtenus n'étaient pas vraiment satisfaisants (notamment par SMA) et qu'ils suggéraient donc que ce type de classification n'est pas possible. Toutefois, ces cinq classes ne sont pas assez peuplées (sept en moyenne) pour tirer des conclusions définitives de ces simulations. De plus, certaines thromboses associées à des cancers (et donc étiquetées *Cancer*) présentent un deuxième facteur de risque (chirurgie ou immobilisation) et auraient donc pu être étiquetées dans une autre classe.

Pour éviter la confusion due aux thromboses multifactorielles, nous avons donc simplifié nos classes en distinguant les thromboses provoquées (*i.e.* avec au moins un facteur de risque connu) des thromboses idiopathiques. Ces deux groupes de thromboses ont ensuite été séparés en deux sous-ensembles en fonction de la survenue d'une embolie pulmonaire. Cette classification en quatre classes de la base de données a donné de meilleurs résultats (notamment en non supervisé) que celle en cinq classes évoquée au paragraphe précédent. Cependant, les résultats ne sont pas encore assez satisfaisants (un peu moins de 50 %) mais les matrices de confusion de certaines simulations se sont montrées intéressantes : la confusion se situait principalement sur la présence ou l'absence d'embolie pulmonaire mais les thromboses provoquées et idiopathiques semblaient plutôt bien séparées. Cette tendance s'est confirmée lorsque nous avons réalisé les simulations en deux classes : thromboses provoquées ou idiopathiques. En effet, dans les cas les plus favorables, les simulations sur les échographies ont atteint les 71 % de bonne classification, ce qui est plutôt encourageant. Les meilleures performances ont été obtenues avec le logarithme des *scattering* coefficients, avec les multicorrélations d'ordre deux ou avec la combinaison de ces deux types de descripteurs. Dans de futurs travaux, il sera donc intéressant de vérifier cette tendance sur une base plus importante. En revanche, les simulations sur les élastographies ont montré une plus grande confusion entre ses deux classes.

Ensuite, les images ont été étiquetées en deux classes en fonction de la survenue d'une embolie pulmonaire. Les résultats se sont avérés encore plus encourageants puisque les simulations les plus favorables sur les échographies ont permis d'obtenir jusqu'à 76 % de bonne classification. Les meilleures performances ont été obtenues sur les descripteurs multirésolutions (sans logarithme) avec les k -moyennes, sur les descripteurs statistiques d'ordre deux avec la classification spectrale ou sur la combinaison des deux types de descripteurs (quel que soit l'algorithme de classification). Les résultats sur les élastographies sont également encourageants (jusqu'à 66 %), notamment en supervisé, mais restent en-dessous de ceux obtenus sur les échographies. Il sera néanmoins intéressant de réaliser des simulations similaires lorsque nous disposerons de suffisamment de données avec le nouveau système Toshiba Aplio 800 (disponible depuis juin 2018 par l'Unité d'Echo-doppler et de Médecine Vasculaire du CHU de Brest). En effet, ce système, plus récent que le système Toshiba Aplio 500, permet d'améliorer significativement la qualité des élastographies.

L'annexe G propose des simulations complémentaires à celles présentées dans cette section. La classification des thromboses par rapport à leur association à un cancer a donné des performances plutôt faibles. Cela pourrait signifier que la présence d'un cancer n'a pas une grande influence sur l'échogénicité ou la dureté du thrombus. De manière similaire, l'étude de l'impact du tabagisme sur la structure du thrombus (classification en deux classes suivant que les patients sont fumeurs ou non-fumeurs) ne s'est pas avérée beaucoup plus performante bien que certaines simulations favorables ont donné quasiment 64 % de bonne classification. Enfin, l'annexe G.3 envisage également l'impact des paramètres d'acquisition des échographies sur les résultats de classification. Les résultats de ces simulations ont montré que, la plupart du temps, le changement de mode d'échographie (fondamental

ou harmonique) et l'ajustement du gain ne semblent pas impacter outre mesure la classification des images. Les simulations sur les glandes salivaires tirent également dans ce sens puisque nous avons obtenu des résultats prometteurs malgré les changements de paramètres entre les différentes acquisitions. Pour mesurer plus précisément cet impact, nous pourrions envisager de construire une base de données dans laquelle les acquisitions de chaque patient seraient collectées pour différents jeux de paramètres prédéfinis.

5.4 Conclusion sur la partie simulation

Cette section a présenté une grande partie des simulations réalisées durant la thèse. Dans un premier temps, nous avons cherché à valider l'implémentation des approches considérées (*Scattering Operator* (SO) et Statistiques d'Ordre Supérieur (SOS)) et à sélectionner les paramètres les plus susceptibles de rendre ces approches efficaces sur nos images de thrombi. Ces deux étapes de validation et d'optimisation ont été effectuées sur la base de référence contenant dix catégories d'images texturées. La classification supervisée (par analyse en composantes principales) a donné d'excellents résultats : jusqu'à 95 % pour les SOS et jusqu'à 99 % pour le SO ou les types de descripteurs combinés (avec 50 % des images consacrées à l'apprentissage). En diminuant la proportion d'images dans la base d'apprentissage, les performances ont logiquement diminué (d'environ 20 %) mais restent très bonnes. La classification non supervisée a abouti sur une nouvelle baisse des résultats : entre 40 et 60 %. Globalement, la classification spectrale a surpassé l'algorithme des k -moyennes (environ 5 % d'écart). Ces taux restent satisfaisants au vu du nombre de classes (dix) et au vu de la proximité de certaines d'entre elles (classes *Papier de verre Eponge* par exemple). De plus, il a été intéressant de noter que l'utilisation conjointe des descripteurs multirésolutions (SO) et statistiques (SOS) a permis d'améliorer la discrimination des dix classes (gain de 3 à 5 %).

Toutefois, nous pouvons remarquer que la base de référence est assez éloignée, en termes de type d'images, de nos données de thrombi veineux. Dans un second temps, nous avons testé nos approches, paramétrées à partir des résultats obtenus sur la base de référence, avec une base de glandes salivaires. Cette base présente le double avantage d'être constituée d'échographies et d'élastographies acquises avec le système Toshiba Aplio 500 (même types d'images que les thrombi) et d'être *a priori* plus simple en termes de classification. L'objectif est de détecter les glandes atteintes du syndrome de Gougerot-Sjörden (classification binaire). Des études médicales récentes montrent que cette maladie est détectable par échographie avec une sensibilité moyenne de 71 % et une spécificité de 92 %. Nos simulations sont donc prometteuses puisque nos approches nous ont permis d'atteindre une sensibilité de 80 % (supervisé ou non supervisé) et une spécificité de 79 % en supervisé (contre 70 % en non supervisé). En considérant uniquement les glandes parotides, la sensibilité et la spécificité atteignent 85 %. Nos résultats sur les élastographies sont également cohérents avec ceux de la littérature médicale avec des sensibilités (85 % par k -moyennes et 81 % pour [83]) et des spécificités (65 % par k -moyennes et 67 % pour [83]) élevées pour les parotides et faibles pour les submandibulaires. En comparaison, le SO (avec logarithme) et les multicorrélations (d'ordre deux) ont donné des résultats plutôt équivalents mais les meilleurs d'entre eux n'ont pas été obtenus avec les mêmes paramètres de simulation. L'association des deux types de descripteurs a souvent permis d'améliorer les performances de classification (ou au moins donné des résultats équivalents). Hormis pour les élastographies des parotides, la classification spectrale a bien souvent surpassé l'algorithme des k -moyennes, comme pour la base de référence.

L'évaluation de nos approches sur les glandes salivaires s'étant avérée concluante, nous nous sommes ensuite consacrés à la classification des thrombi. Si le syndrome de Gougerot-Sjörden est bien souvent détectable par échographie, il n'existe actuellement aucune étude qui montre qu'il est possible d'identifier l'origine de la formation d'un thrombus humain ou d'estimer le risque de survenue d'une embolie pulmonaire par échographie et élastographie. Nos premières simulations ont montré qu'il était peut-être optimiste de vouloir retrouver le facteur déclencheur principal (cancer, immobilisation, chirurgie, thrombophilie, idiopathique) à partir de ces deux types d'images : ce type de classification est

d'autant plus compliqué que l'apparition d'une Thrombose Veineuse Profonde (TVP) provoquée (avec un facteur déclencheur connu et identifiable) est souvent multi-factorielle. Nous avons alors simplifié nos classes en considérant uniquement la présence d'un facteur déclencheur connu (provoquée ou idiopathique) et celle d'une embolie pulmonaire. En effet, ces deux conditions jouent un rôle crucial dans le choix et la durée du traitement de la thrombose. Par exemple, une thrombose provoquée est bien souvent suivie d'un traitement anticoagulant plus long que pour une thrombose idiopathique. Les résultats de classification en quatre classes nous ont encouragés à dissocier la classification de ces deux conditions. Nous avons alors obtenu des résultats de classification en deux classes très encourageants : nos approches ont permis d'identifier à 70 % le caractère provoqué ou idiopathique de la thrombose et de détecter à 76 % la présence d'embolie pulmonaire. Ces résultats ont été obtenus sur les échographies mais l'élastographie a également donné des résultats intéressants pour détecter la présence d'embolie pulmonaire (66 %). Globalement, l'association des deux types de descripteurs a, une nouvelle fois, permis de faire aussi bien, voire mieux, que leur utilisation séparée. Les simulations présentées dans ce mémoire ont toujours considéré de manière disjointe l'échographie et l'élastographie. A l'avenir, il sera donc très intéressant de combiner les descripteurs obtenus sur ces deux types d'image pour vérifier si cela peut permettre d'améliorer la classification des thrombi (et des glandes salivaires).

Conclusion générale et perspectives

Conclusion générale

Ces travaux de thèse se situent dans le contexte particulier du traitement d'images biomédicales. L'objectif est de développer une chaîne de traitement pour identifier et caractériser un thrombus veineux par imagerie acoustique. La formation inappropriée d'un thrombus dans une ou plusieurs veines du réseau sanguin profond (par opposition à superficiel) conduit au développement d'une thrombose veineuse profonde, plus connue sous le nom de phlébite. Le principal danger d'une thrombose veineuse profonde est qu'un fragment du thrombus (ou sa totalité) se détache de la paroi veineuse, remonte vers le cœur en étant entraîné par la circulation sanguine puis se loge dans une artère pulmonaire. Cette complication, appelée embolie pulmonaire, est la troisième cause de mortalité vasculaire après l'infarctus du myocarde et l'accident vasculaire cérébral.

La thrombose veineuse profonde est une maladie multifactorielle associant des facteurs génétiques et acquis. Trois mécanismes physiopathologiques interviennent, de manière isolée ou associée, dans son développement : une stase veineuse (*i.e.* ralentissement de la circulation), une altération de la paroi veineuse et une thrombophilie (*i.e.* un renforcement de la coagulation ou un déficit de l'anticoagulation). Parmi les facteurs de risque les plus fréquents, un cancer, un acte de chirurgie (anesthésie, saignement, pose d'un cathéter), une immobilisation (plâtre, hospitalisation, position assise prolongée) ou une grossesse peut déclencher un ou plusieurs de ces trois mécanismes et déboucher sur une thrombose. Toutefois, il est courant qu'une thrombose se manifeste sans raison apparente ; elle est alors appelée idiopathique. Notre projet concerne l'analyse de la structure de thrombi veineux humain dans le but de prédire le risque d'embolie pulmonaire et/ou d'identifier l'origine de la thrombose. Il est connu qu'un thrombus contient des globules rouges et des plaquettes dans un réseau de fibrine dont la stabilité et la microstructure sont dépendantes de l'ensemble des processus biologiques qui ont abouti à sa formation. L'enjeu de notre projet a donc été d'observer (par imagerie acoustique) si la structure du thrombus est différente suivant le mécanisme qui l'a généré et suivant la survenue d'une embolie pulmonaire.

Depuis plusieurs décennies, l'échographie est le mode d'imagerie le plus utilisé pour diagnostiquer une thrombose veineuse profonde : en présence d'un thrombus, la veine est difficilement, voire pas du tout, compressible. De plus, la région de la veine où se situe le thrombus est généralement plus échogène que l'artère voisine. Si la phlébographie (radio des veines avec un produit de contraste) reste encore aujourd'hui l'examen de référence, l'échographie présente l'avantage d'être efficace (excellente sensibilité), quasiment sans danger (indolore et non irradiante), mobile (possibilité de réaliser l'examen dans la chambre du patient), en temps réel et peu coûteuse. Plus récemment, une nouvelle technique d'imagerie acoustique, l'élastographie, fait son apparition dans de nombreux systèmes d'échographie dont le système Toshiba Aplio 500 utilisé pour collecter nos données. L'élastographie analyse la propagation des ondes de cisaillement (l'échographie exploite les ondes de compression) dans les tissus humains pour estimer qualitativement et/ou quantitativement la dureté des tissus examinés. Cette technique d'imagerie, qui a d'abord montré son efficacité pour le diagnostic d'une fibrose hépatique, présente un grand nombre d'applications : sein, thyroïde, prostate, reins, *etc.* Parmi les applications récentes de l'élastographie, il y a celle des glandes salivaires et plus particulièrement le diagnostic du syndrome de Gougerot-Sjörden (maladie auto-immune qui se manifeste par une sécheresse buccale et oculaire). Cette application est d'autant plus intéressante que le service de Rhumatologie du CHU de Brest a construit une base de données avec le même système d'acquisition que nos images de thrombi.

Pour répondre à notre problématique, l'Unité d'Echo-doppler et de Médecine Vasculaire, disposant d'un système Toshiba Aplio 500, nous a permis de collecter des échographies et des élastographies de thrombi veineux. Au cours des trois années de thèse, la société Toshiba a effectué une mise à jour du système et a fourni une sonde linéaire qui a permis d'améliorer grandement la qualité des élastographies. L'Unité d'Echo-doppler et de Médecine Vasculaire a également développé un bras articulé pour limiter la variabilité intra- et inter-opérateur (la force de pression de la sonde sur la peau du patient peut modifier l'élastographie). Toutefois, nous nous sommes aperçus que la variabilité intra-opérateur (*i.e.* même opérateur, même patient, même zone examinée) des élastographies obtenues avec la sonde linéaire était encore assez importante. Le retour à la sonde convexe a permis de stabiliser les acquisitions d'élastographie. De plus, l'amélioration de la qualité des élastographies par la mise à jour de l'équipement a permis de réduire le nombre d'acquisitions par patients (de dix à trois acquisitions). Ce gain de temps a permis au praticien d'acquérir des images en coupe longitudinale en plus des acquisitions en coupe transversale. Ces évolutions du protocole d'acquisition ont débouché sur la création de trois bases distinctes (*cf.* chapitre 1). Ce mémoire présente principalement les travaux réalisés sur la dernière base de données.

L'objectif de ces travaux de thèse était donc de préparer et d'analyser les images acoustiques de thrombi pour caractériser leur structure et retrouver, seulement à partir des images, le facteur responsable de leur formation et la présence d'embolie pulmonaire. Le chapitre 2 propose une revue des travaux réalisés pour caractériser le thrombus veineux par échographie et/ou élastographie. Nous avons pu constater qu'il y a, actuellement, relativement peu de travaux sur notre problématique et que les études existantes sont bien souvent réalisées à partir de thrombi animaux (souvent *in vitro*) dans le but de les dater. Par conséquent, l'état de l'art a été étendu aux méthodes existantes dans le domaine biomédical et, de manière plus générale, à celles du traitement d'images. Ce chapitre 2 présente également les classifieurs qui ont été utilisés pour nos simulations (sélection de modèles affines, k -moyennes et classification spectrale).

Notre base de données est constituée de captures d'écran du système Toshiba Aplio 500 stockées au format DICOM (*Digital Imaging and Communications in Medicine*). La première partie des travaux de thèse a été consacrée à la préparation des données (*cf.* chapitre 3). En effet, le but est, ici, de récupérer les échographies et les élastographies présentes dans ces captures mais également l'ensemble des paramètres d'acquisition, des données relatives aux patients et à la formation des thrombi. Ces informations sont enregistrées soit dans les métadonnées du fichier DICOM, soit directement sur la capture d'écran, soit dans un tableur (pour les informations médicales). J'ai donc réalisé une première Interface Homme-Machine (IHM) qui permet de réaliser cette tâche. Les images et les données extraites d'une même acquisition sont alors stockées dans une même structure. A ce niveau-là, les images contiennent le thrombus mais aussi toutes les structures avoisinantes. Initialement, j'avais proposé de segmenter l'ellipse dessinée autour du thrombus par le praticien pour extraire une imagelette contenant du thrombus. Cette méthode possédait l'inconvénient de devoir redimensionner l'imagelette extraite et donc de modifier l'échelle de l'image. Par conséquent, j'ai développé une seconde IHM pour segmenter manuellement les échographies de thrombi. Ensuite, à partir de l'image segmentée, trois types d'imagelettes sont extraites automatiquement : celles contenant exclusivement du thrombus, celles contenant une grande partie de thrombus (mais aussi de la paroi et des tissus voisins) et celles ne contenant pas de thrombus. L'intérêt d'extraire des imagelettes ne contenant pas de thrombus est de pouvoir prendre en compte le contraste entre le thrombus et son voisinage lors de la classification. Pour les élastographies, j'ai préféré créer une troisième IHM qui permet de sélectionner manuellement les imagelettes qui seront extraites : ces cartes de dureté possèdent des régions sans valeur (zones non atteintes par les ondes de cisaillement) qui pourraient mettre en péril l'extraction automatique d'imagelettes.

Après avoir préparé les données de thrombi, nous disposons donc d'un certain nombre d'images pour chaque acquisition. Nous pourrions appliquer directement des outils de classification sur les images pour observer si les thrombi possèdent des structures différentes. Toutefois, la littérature montre qu'il est souvent préférable de transformer les images pour obtenir le résumé le plus pertinent possible des images. Cette étape, appelée *extraction de descripteurs*, constitue une partie importante des travaux de thèse. Dans ce mémoire, nous proposons deux approches distinctes pour extraire des descripteurs de nos images : le *scattering operator* et les statistiques d'ordre supérieur. Le *scattering operator* est un outil puissant qui utilise les transformées en ondelettes pour extraire des caractéristiques multirésolutions des images. Il est particulièrement adapté pour discriminer des textures et a notamment montré son efficacité en classification des fonds marins par imagerie acoustique (sonar). De nombreuses techniques du traitement d'images (comme l'analyse en composantes principales) sont basées sur les statistiques d'ordre deux (écart-type, variance, covariance, corrélation). Toutefois, il a été montré dans de nombreuses applications (parole, électrocardiogramme) que l'utilisation des statistiques d'ordre supérieur à deux (l'ordre quatre notamment) permettait d'améliorer et d'étendre les solutions existantes à l'ordre deux, voire de résoudre des problèmes non accessibles à l'ordre deux. Ces deux approches présentent également l'avantage d'être complémentaires puisqu'elles n'exploitent pas la même source d'information issue des images.

Chaque image de la base de données a été représentée par des descripteurs multirésolutions (obtenus en appliquant le *scattering operator*) et des descripteurs statistiques : moyenne (ordre un), autocorrélation (ordre deux), bicorrélation (ordre trois) et tricorrélation (ordre quatre). La dernière étape de la chaîne de traitement d'images consiste à classer les descripteurs extraits des images. Pour mener cette tâche, nous avons considéré quatre algorithmes de classification parmi les plus usuels :

- En mode supervisé (avec une partie des images consacrées à l'apprentissage) :
 1. La sélection de modèles affines ;
- En mode non supervisé :
 2. Les k -moyennes ;
 3. Le DBSCAN ;
 4. La classification spectrale.

Comme nous pouvons le constater, nous avons privilégié la piste non supervisée car nous ne disposons pas de données en quantité suffisante (trente-sept patients) pour certaines techniques de la classification supervisée comme les machines à vecteurs de support ou les réseaux de neurones.

Les approches d'extraction de descripteurs et de classification ont d'abord été validées et optimisées sur une base de référence (base contenant des images synthétiques de textures) avec succès. Elles ont ensuite été évaluées sur une base plus proche de celle des thrombi : la base de glandes salivaires. Les simulations sur les échographies de glandes salivaires ont donné des résultats proches de ceux obtenus dans des études médicales (diagnostic du syndrome de Gougerot-Sjörden par des experts médicaux). En classification supervisée, nous avons atteint une sensibilité de 80 %, supérieure aux études médicales récentes (71 % en moyenne), et une spécificité de 79 %, inférieure à ces mêmes études (92 % en moyenne). En mode non supervisé, la classification spectrale a permis d'atteindre une sensibilité identique de 80 % et une spécificité moindre de 71 %. Les simulations sur les élastographies ont donné des résultats de classification plus faibles car il est plus compliqué de détecter le syndrome de Gougerot-Sjörden par élastographie sur les submandibulaires. En revanche, en ne considérant que les glandes parotides, nous parvenons à des résultats similaires à ceux obtenus par une étude médicale : une sensibilité de 85 % par les k -moyennes (81 % pour l'étude) et une spécificité de 65 % par k -moyennes (67 % pour l'étude). Il est intéressant de noter que l'utilisation conjointe des descripteurs multirésolutions et statistiques permet en général d'améliorer la détection du syndrome de Gougerot-Sjörden.

Après avoir obtenu des résultats concluants sur notre base de référence et sur la base des glandes salivaires, j'ai lancé les simulations sur la base des thrombi. J'ai d'abord voulu retrouver les facteurs déclencheurs des thrombi (cancer, immobilisation, chirurgie, thrombophilie et idiopathique) mais nous avons observé une grande confusion entre les cinq classes proposées. Toutefois, ce résultat n'est pas illogique car certains patients présentent plusieurs facteurs de risque (par exemple cancer et immobilisation). De plus, si les thromboses par immobilisation sont généralement dues à une stase veineuse (et peut-être une thrombophilie), les cancers ou la chirurgie peuvent actionner plusieurs mécanismes responsables de la coagulation. Il est donc plus pertinent d'essayer de classer les thrombi en fonction de deux autres paramètres : le caractère provoqué ou idiopathique de la thrombose, et la survenue ou non d'une embolie pulmonaire. Ces deux paramètres sont d'autant plus pertinents qu'ils représentent les principaux critères dans le choix du traitement de la thrombose. La classification en quatre classes s'est avérée meilleure que celle en fonction des facteurs déclencheurs mais pas encore suffisamment satisfaisante. En revanche, en considérant la classification de ces deux paramètres séparément, nous avons obtenu des résultats beaucoup plus intéressants. En élastographies, les performances de classification ont atteint 66 % de bon classement des thromboses en fonction de la présence d'embolie pulmonaire. En échographie, les simulations ont permis de séparer correctement les thromboses provoquées des thromboses idiopathiques à hauteur de 70 % et de détecter la présence d'une embolie pulmonaire à hauteur de 76 %.

En résumé, au cours de ces travaux de thèse, nous avons adapté un ensemble d'outils du traitement du signal et du traitement d'images à des données médicales : des échographies et des élastographies. Nous nous sommes consacrés à chaque étape de la chaîne de traitement d'images avec une attention particulière sur l'étape d'extraction de caractéristiques. Cette chaîne a d'abord été validée sur des images synthétiques de textures puis sur des données médicales de même nature (application des glandes salivaires). Pour caractériser la structure des thrombi veineux et identifier si cette dernière diffère en fonction des facteurs déclencheurs et de la présence (ou l'absence) d'embolie pulmonaire associée, nous avons mis en place un protocole expérimental afin d'analyser les performances de classification pour différents paramétrages de la chaîne de traitement (descripteurs, pris en compte du voisinage du thrombus, *etc*). Les résultats de classification obtenus sont prometteurs, notamment pour la prédiction d'embolie pulmonaire à partir des images de thrombi. Par conséquent, ces travaux novateurs (application d'outils du traitement du signal pour la caractérisation du thrombus veineux) ouvrent de nombreuses perspectives à la fois médicales, applicatives et dans le traitement des images.

Perspectives

D'un point de vue applicatif, nous projetons de développer une IHM qui permettrait, à chaque nouvelle acquisition réalisée par l'Unité, de préparer les images collectées, d'en extraire les imagerie, de calculer les descripteurs et de mettre à jour les résultats de classification (de manière automatique). Cette IHM permettrait d'agrandir progressivement la base de données et devrait, en outre, améliorer la modélisation des classes recherchées en mode supervisé.

L'Unité d'Echo-doppler et de Médecine Vasculaire dispose depuis juin 2018 du dernier système de Toshiba (l'Aplio 800) qui permet d'améliorer la qualité des élastographies de manière significative et de générer des images en trois dimensions du thrombus. De plus, l'Unité s'est également munie d'un fantôme de test (Gammex Sono TE Phantom) qui pourra nous aider à standardiser encore davantage nos acquisitions. En effet, l'utilisation de ce fantôme au cours d'un examen va nous permettre de calibrer l'échographie et l'élastographie et de favoriser l'uniformité de nos acquisitions. Plus globalement, le CHRU de Brest est en train de mettre au point un protocole de recherche clinique pour la collecte des données de thromboses veineuses qui inclurait, en plus des images échographiques et élastographies, des données biologiques (comme les D-dimères) et un suivi régulier du patient (un examen par semaine). La construction d'une nouvelle base avec le nouveau système et ce protocole va nous permettre de mieux caractériser la structure du thrombus et de confirmer les résultats obtenus sur la base considérée dans ce mémoire.

Au niveau de la chaîne de traitement d'images, nous pourrions d'abord procéder à une fusion (*i.e.* l'association) des descripteurs issus des échographies avec ceux issus des élastographies pour classer les images de thrombi (et de glandes salivaires). Cette fusion des informations issues des deux modes d'imagerie permettrait certainement d'augmenter les performances de classification : il faudrait probablement donner des poids différents aux deux modes d'imagerie. Il serait également intéressant de classer nos descripteurs avec de nouveaux classifieurs (comme les machines à vecteurs de support) et d'utiliser l'astuce du noyau sur les algorithmes proposés dans ce mémoire : cette astuce consiste à projeter les données dans un espace de plus grande dimension pour que des frontières, non linéaires dans l'espace de départ, deviennent linéaires dans cet espace plus vaste. Au niveau des statistiques d'ordre supérieur, nous avons considéré seulement certains axes des tenseurs des multicorrélations. L'utilisation de techniques de contraction de tenseurs permettrait probablement de mieux exploiter l'information présente dans les multicorrélations. De plus, nous nous sommes intéressés aux statistiques d'ordre supérieur dans le domaine spatial mais il est également possible de les définir dans le domaine fréquentiel : les multispectres. De manière similaire, une meilleure fusion des descripteurs statistiques et multirésolutions permettrait de mieux exploiter l'information pertinente de ces descripteurs et ainsi d'améliorer leur classification. Enfin, nous pourrions enrichir chaque étape de la chaîne de traitement par de nouvelles méthodes issues de la littérature et d'évaluer leur impact sur les résultats de classification.

Liste des publications

A. Article de journaux publié

Advances in Science, Technology and Engineering Systems Journal (ASTESJ)

A.1. Deep venous thrombus characterization : ultrasonography , elastography and scattering operator

Cet article de journal [11] est une suite des deux articles de conférence présentés à ICFSP 2016. Il présente la problématique (médicale et technique) de notre projet, le *scattering operator* et son application sur la base collectée avec le protocole 1. Les coefficients du *scattering operator* de chaque image sont ensuite projetés dans un plan par ACP pour étudier leur dispersion en fonction des paramètres physiopathologiques des patients (facteurs déclencheurs et présence EP).

Auteurs : T. Berthomier, A. Mansour, L. Bressollette, F. Le Roy, and D. Mottier

Abstract : *A thrombus or a blood clot is the result of blood coagulation which is a natural process to prevent bleeding. An inappropriate formation of a thrombus in a deep vein is known as Deep Venous Thrombosis (DVT). The main complication of a DVT is a Pulmonary Embolism (PE) which occurs when a thrombus breaks loose and travels to the lungs. DVT, PE, or both are also known as Venous thromboembolism (VTE). It affects an estimated 300,000–600,000 Individuals just in the United States per year and can cause considerable morbidity and mortality. This multifactorial disease related to advanced age, immobility, surgery or obesity is an important public health issue. Our project is looking to link the VTE epidemiology (risk factors, patient history, PE) to the thrombus structure. To reach our goals, we are collecting ultrasonography (echogenicity) and elastography (stiffness) of human thrombus. This manuscript describes our approach to create and preprocess a database using Toshiba Aplio 500. Our approaches to characterize the thrombus structure with ultrasound images are also described. The feature extraction is made with the scattering operator. Obtained features are then reduced using Principal Component Analysis and are analyzed to evaluate our approach.*

B. Articles de conférences publiés

International conference on Bioinformatics and Biomedicine (BIBM)

B.1. Unsupervised clustering of DVT Ultrasound Images using High Order Statistics

Cette publication a été présentée lors du Workshop MABM (Machine Learning and Artificial Intelligence in Bioinformatics and Medical Informatics) de la conférence BIBM qui s'est déroulée du 3 au 6 décembre 2018 à Madrid (Espagne). Dans cet article [13], nous envisageons les statistiques d'ordre supérieur pour extraire des descripteurs de nos images de thrombi. Cette approche statistique est d'abord appliquée sur une base de test KTH-TIPS, puis sur la base collectée avec le protocole 2B (afin de comparer les résultats obtenus avec ceux du *scattering operator* présentés à la conférence ICBSIA) et enfin sur la base collectée avec le protocole 3 (cf. 1.3.2.3).

Auteurs : T. Berthomier, A. Mansour, L. Bressollette, D. Mottier, F. Le Roy, B. Hermenault, C. Hoffmann, and L. Frechier

Abstract : *Naturally, a thrombus (or a blood clot) is developed by our body to prevent bleeding. However, various factors can slow or change the blood flow and create a thrombus in a deep vein (more often in the legs). This inappropriate situation is called Deep Venous Thrombosis (DVT) and it may permanently damage the blood vessels. Moreover, this disease (DVT) becomes deadly when a broken fragment of the thrombus reaches the lung and generates a Pulmonary Embolism. Using ultrasound images, our project aims to analyse the thrombus structure and extract information leading to its triggering factors : Immobilization, cancer, surgery, genetic variations, pregnancy, age, etc. In previous studies, we developed an approach based on wavelet transform to analyse the ultrasound images : Ultrasonography and elastography. In this manuscript, we extract new features based on High Order Statistics. These statistics provide hidden relevant information in the images to characterize the thrombus structure.*

International Conference on Biomedical Signal and Image Analysis (ICBSIA)

B.2. Scattering operator and spectral clustering for ultrasound images : application on deep venous thrombi

Cette conférence s'est déroulée les 20 et 21 novembre 2017 à Paris. Dans cet article [12], nous proposons de classer les images de thrombi collectées avec le protocole 2B (*cf.* section 1.3.2.2) par classification spectrale. Cette technique est appliquée directement sur les images ou sur leurs descripteurs multirésolutions (*i.e.* calculés par le *scattering operator*). Pour valider nos deux approches et définir au mieux leurs paramètres, le *scattering operator* et la classification spectrale sont appliqués sur une base de test, la base KTH-TIPS [56].

Auteurs : T. Berthomier, A. Mansour, L. Bressollette, F. Le Roy, D. Mottier, L. Frechier, and B. Hermenault

Abstract : *Deep Venous Thrombosis (DVT) occurs when a thrombus is formed within a deep vein, usually in the legs. This situation can be deadly if a part or the whole thrombus reaches the lung and causes a Pulmonary Embolism (PE). This disease, often asymptomatic, has multifactorial causes : immobilization, surgery, pregnancy, age, cancers, and genetic variations. Our project is aiming to relate the thrombus epidemiology (origins, patient predispositions, PE) with its structure using ultrasound images. Ultrasonography and elastography were collected using Toshiba Aplio 500 at Brest Hospital. This manuscript compares two classification approaches to extract discriminative features : spectral clustering and scattering operator. The former is based on the graph and matrix theories while the latter cascades wavelet convolutions with nonlinear modulus and averaging operators.*

International Conference on Frontiers of Signal Processing (ICFSP) :

Cette conférence internationale s'est déroulée du 15 au 17 octobre 2016 à Varsovie (Pologne). Lors de celle-ci, j'ai eu l'occasion de présenter les deux publications suivantes :

B.3. Venous blood clot structure characterization using scattering operator

Cet article [10] décrit le contexte et la problématique de notre projet sur les thromboses veineuses profondes. Ensuite, il présente le *scattering operator* et les premiers résultats obtenus sur la base collectée avec le protocole 1 (*cf.* section 1.3.2.1).

Auteurs : T. Berthomier, A. Mansour, L. Bressollette, F. Le Roy, and D. Mottier

Abstract : *Deep venous thrombosis (DVT) occurs when a blood clot appear within a deep vein, usually in the legs. The main complication is pulmonary embolism (PE) which is the third cause of vascular death after myocardial infarction and cardiovascular event. DVT onset is multifactorial (immobilization, surgery, age, cancers, genetic variations) and it is mostly diagnosed via ultrasound. In our project, we are interested in a new approach, which consists in using ultrasound and elastography to assess the mechanical properties of blood clots and to identify the thrombosis triggering factors. This means characterizing its structure, establishing his age, the cause of its formation and the risk of PE. In this paper, we aim to analyze the clot texture using the scattering operator which combines wavelet transform convolutions with non-linear modulus and averaging operators. The scattering operator is showing promising results in signal processing and especially in image classification. Therefore, we will apply this operator to our database and discuss the results of our simulation at the end of this manuscript.*

B.4. Deep Venous Thrombosis : Database creation and image preprocessing

Cet article [9] présente le système d’acquisition (Toshiba Aplio 500), le protocole de collecte des données (protocole 1 de la section 1.3.2.1), les deux modes d’imagerie (échographie et élastographie) et l’ensemble des étapes de préparation des images (*cf.* annexe B).

Auteurs : T. Berthomier, A. Mansour, L. Bressollette, F. Le Roy, and D. Mottier

Abstract : *Deep Venous Thrombosis (DVT) is a major national health problem. Its progression to a pulmonary embolism can lead to significant mortality. DVT is a frequent, severe and often under-recognized diagnosis for critically ill patient. Many factors can increase the risk of DVT : immobilization, cancer, age, surgery, previous DVT, hormone, thrombophilia, etc. Ultrasonography is mostly performed for its diagnosis because it is easy, immediate and non-invasive. In our project, the main goal is to determine the possible correlations between the blood clot structure and the DVT triggering factors. To characterize its structure, we decided to combine ultrasound and elastography imaging. The latter technique is already used to diagnose other diseases such as liver fibrosis and gives promising results to date animal blood clot. This paper explains how we create our database using Toshiba’s system. Our data are screenshot images that have to be prepared for further processing such as texture characterization.*

C. Articles de journaux en préparation

Nous projetons de soumettre deux articles de journaux. Le premier papier présente l’intérêt de combiner les descripteurs multirésolutions (obtenus avec le *scattering operator*) et les descripteurs statistiques pour caractériser la structure d’un thrombus veineux par échographie et élastographie. Le second manuscrit s’intéresse à l’utilisation conjointe de l’échographie et de l’élastographie des glandes salivaires pour détecter de manière automatique le syndrome de Gougerot-Sjörden. Nous envisageons également de réaliser deux études cliniques (une pour chacune des applications) pour présenter nos résultats dans des revues médicales.

C.1. Scattering and high order statistic features for the characterization of deep venous thrombus : application on ultrasonography and elastography images

C.2. Automatic detection of the Sjögren syndrome detection using ultrasonography and elastography images

Annexe A

Thrombus veineux et Imagerie

Introduction

Cette annexe décrit les processus naturels qui participent à la création puis à la destruction d'un thrombus [77] puis elle énumère les principaux outils utilisés pour déceler la présence d'un thrombus dans le réseau veineux et l'analyser *in vivo*. La section A.3 est consacrée aux différents phénomènes physiques qui se produisent lorsqu'une onde acoustique se propage dans un milieu inhomogène comme le corps humain.

A.1 Hémostase et fibrinolyse

L'hémostase désigne l'ensemble des mécanismes physiologiques qui assurent le maintien du sang à l'intérieur des vaisseaux et permet notamment l'arrêt du saignement lorsqu'un vaisseau est blessé. L'ensemble des étapes de l'hémostase est résumé sur la figure A.1. L'objectif principal de l'hémostase est de « boucher » la lésion du vaisseau. Ce sont principalement les plaquettes (petits éléments du sang) et le fibrinogène qui vont entrer en action afin de « colmater » la brèche en formant le clou plaquettaire (voir figure A.2).

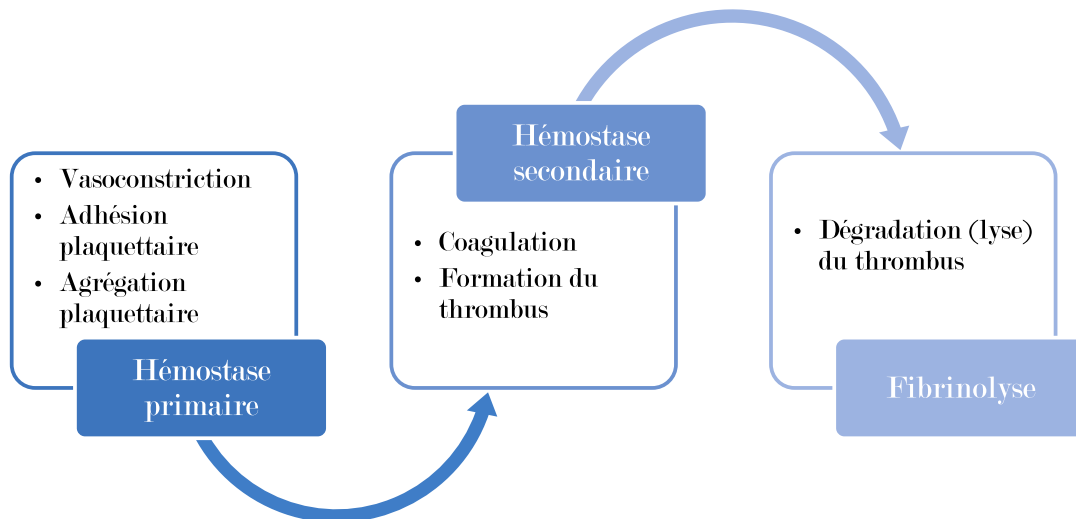


FIGURE A.1 – Ensemble des mécanismes permettant l'arrêt du saignement suite à la lésion d'un vaisseau.

La vasoconstriction est la réponse immédiate à la lésion d'un vaisseau. Il s'agit d'un spasme vasculaire qui diminue le diamètre du vaisseau et ralentit le saignement. Ensuite, les plaquettes vont se lier au collagène des parois vasculaires, c'est l'adhésion plaquettaire (voir figure A.2a). Cette étape est suivie de l'agréation plaquettaire (voir figure A.2b), c'est-à-dire l'adhésion des plaquettes entre elles. Ces trois actions successives sont appelées hémostase primaire.

L'hémostase secondaire désigne l'enchaînement des réactions chimiques impliquant divers substrats et enzymes plasmatiques nécessaires à la coagulation. Ces interactions complexes débouchent au bout du compte en la transformation d'une protéine soluble, le fibrinogène, en une protéine insoluble, la fibrine, qui forme l'armature du thrombus.

La fibrinolyse est le processus par lequel la fibrine est détruite (lyse). Elle est indispensable pour dissoudre le thrombus. Ce processus clôture la coagulation sanguine en perméabilisant les vaisseaux sanguins réparés et sert à éviter l'apparition de phlébite. Les troubles de l'hémostase, tels que la thrombopénie (insuffisance de plaquettes) et la carence en facteurs de coagulation, augmentent le risque de phlébite et ils sont considérés à la sous-section suivante.

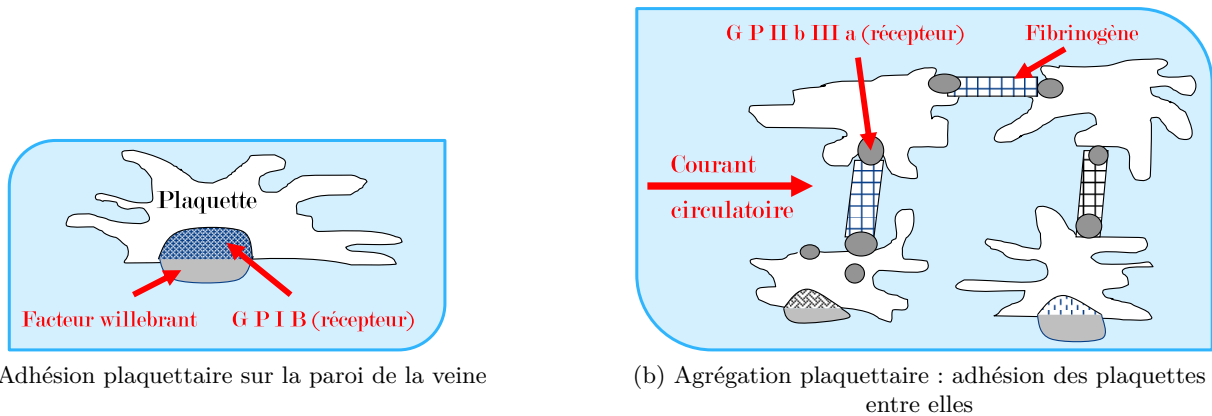


FIGURE A.2 – Clou plaquettaire : adhésion plaquettaire et agrégation plaquettaire.

A.2 Principaux outils de détection d'un thrombus

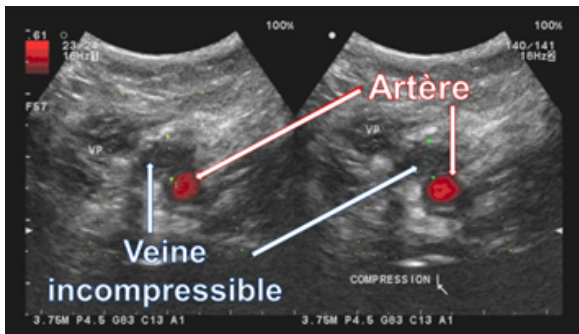
A.2.1 Phlébographie

La phlébographie [95] est un examen radiologique des veines après injection d'un produit de contraste iodé. La quantité de radiation dépend de la profondeur de la zone étudiée et du produit de contraste. Actuellement, cet examen n'est plus utilisé en pratique quotidienne à cause des radiations (et les risques d'allergies qu'elles entraînent) et car le doppler veineux le remplace avantageusement. Toutefois, il reste l'examen de référence dans de nombreuses études cliniques du fait de résultats non opérateur-dépendants et de sa reproductibilité interobservateur.

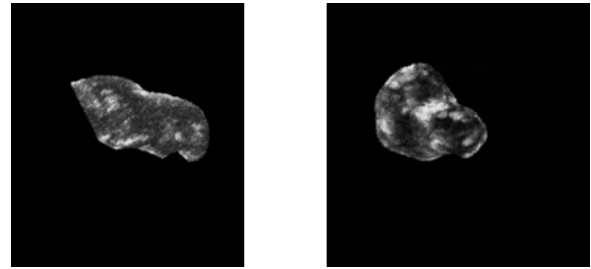
A.2.2 Doppler veineux

Le doppler veineux est l'examen le plus courant pour diagnostiquer une thrombose veineuse. Il utilise la technologie ultrasonore, qui est peu coûteuse, totalement non invasive, indolore et sans radiation, couplée d'un module doppler qui permet de mesurer la vitesse des flux sanguins. L'échographie permet d'obtenir une image (voir figure A.3.b) en niveau de gris dont l'amplitude dépend de l'échogénicité des tissus. L'utilisation du module doppler donne les vitesses des flux sanguins et indique avec un disque bleu ou rouge le sens de circulation du flux dans les artères et les veines (si le sang circule).

Les signes directs de thrombose veineuse sont la visualisation d'un matériel endoluminal plus ou moins échogène qui correspond au thrombus (voir figure A.3.a) et l'absence de flux dans la veine. Le thrombus peut être totalement adhérent à la paroi veineuse ou au contraire flottant dans la lumière. Les signes indirects sont l'incompressibilité veineuse (voir figure A.3.b), la parésie valvulaire (perte de mobilité des valvules à cause du thrombus) et l'augmentation du flux dans les veines collatérales.



(a) Doppler veineux avec un thrombus dans la veine poplitée [49]



(b) Rendu volumique de deux thrombus par échographie [19]

FIGURE A.3 – Echogénéicité d'un thrombus et détection d'un thrombus par doppler veineux.

A.2.3 Angiographie par Résonance Magnétique

L'Angiographie par Résonance Magnétique (AMR), comme l'Imagerie par Résonance Magnétique (IRM), utilise le signal émis par les protons d'hydrogène alignés sous l'effet d'un puissant champ magnétique pour visualiser veines ou artères. En général, cette technique est employée avec un produit de contraste. Il existe quelques études [95] qui comparent les bonnes performances de cet examen avec celles de la phlébographie malgré le nombre limité de données patients ARM.

A.2.4 Angioscanner

Au même titre que les deux examens évoqués précédemment aux sections A.2.1 et A.2.3, l'angioscanner nécessite souvent l'injection d'un produit de contraste et expose le patient à des radiations. Il est ainsi peu utilisé dans la pratique quotidienne pour détecter les thromboses veineuses en comparaison du doppler veineux qui est non invasif. Cet examen est, le plus souvent, utilisé en complément du doppler veineux chez un patient suspecté d'embolie pulmonaire [95].

A.3 Propagation acoustique dans un milieu inhomogène

Dans un milieu homogène et non absorbant, l'onde ultrasonore conserve une amplitude et une direction constante [40, 53]. En pratique, l'onde qui se propage dans les tissus biologiques voit son intensité décroître au fur et à mesure de sa pénétration, sa direction modifiée et sa forme altérée. Plusieurs phénomènes sont à l'origine de ces perturbations :

- **La diffusion** : si la taille de l'inhomogénéité est très inférieure à la longueur d'onde λ , l'énergie acoustique se disperse de façon homogène dans toutes les directions de l'espace.
- **Les interférences** : lorsque l'onde atteint plusieurs obstacles proches les uns des autres (des diffuseurs par exemple), chacun se comporte comme une source ponctuelle et les différentes ondes émises vont former une seule onde dont l'intensité est très variable. Suivant le rapport entre la longueur d'onde et la distance des cibles, les interférences peuvent être destructives ou constructives et sont à l'origine de l'aspect en grains (appelé *speckle* des images échographiques).
- **La diffraction** : lorsqu'une onde se rapproche d'un obstacle à une distance proche de sa longueur d'onde, elle est déviée de sa trajectoire originelle. De même, lorsqu'elle rencontre un obstacle de dimension limitée par rapport à sa longueur d'onde, sa direction de propagation est modifiée.
- **L'absorption** : ce phénomène transforme l'énergie incidente de l'onde ultrasonore en chaleur et constitue une perte importante d'énergie.

• **L'atténuation** : l'atténuation correspond à la diminution de pression ultrasonore par absorption et diffusion. Elle englobe l'ensemble des pertes d'énergie dues à la propagation de l'onde dans le milieu. Cette énergie ne sera donc pas capturée par la sonde réceptrice. L'atténuation des ultrasons décroît globalement de façon exponentielle avec la distance parcourue. Elle peut s'exprimer par un coefficient α donné généralement en $dB \cdot cm^{-1}$:

$$\alpha = -\frac{10}{x} \log \frac{I_x}{I_0} \quad (\text{A.1})$$

où x représente la distance parcourue en cm , I_0 et I_x les intensités de pression acoustiques respectivement émises dans le milieu et transmises en sortie du milieu. Quelques ordres de grandeur de ce coefficient α pour différents milieux sont donnés dans le tableau 1.1. Ce coefficient d'approximation est lui-même proportionnel à la fréquence de l'onde :

$$\alpha [db.cm^{-1}] = \alpha [db.cm^{-1}.MHz^{-1}] f^\beta [MHz] \quad (\text{A.2})$$

où β est compris entre 1 et 2 pour les tissus absorbants. Pour l'eau, β vaut 2. En général, β est fixé à 1 pour les tissus mous. L'atténuation est donc beaucoup plus élevée pour les hautes fréquences d'où l'utilisation de fréquences basses pour l'exploration d'organes ou de tissus plus profonds.

Annexe B

Evolution de la chaîne de prétraitements

Glossaire de l'annexe B

BDD Base De Données

DICOM Format standard pour la gestion des images biomédicales (*Digital Imaging and Communications in Medicine*)

IHM Interface Homme-Machine

TCD Transformée en Cosinus Discrète

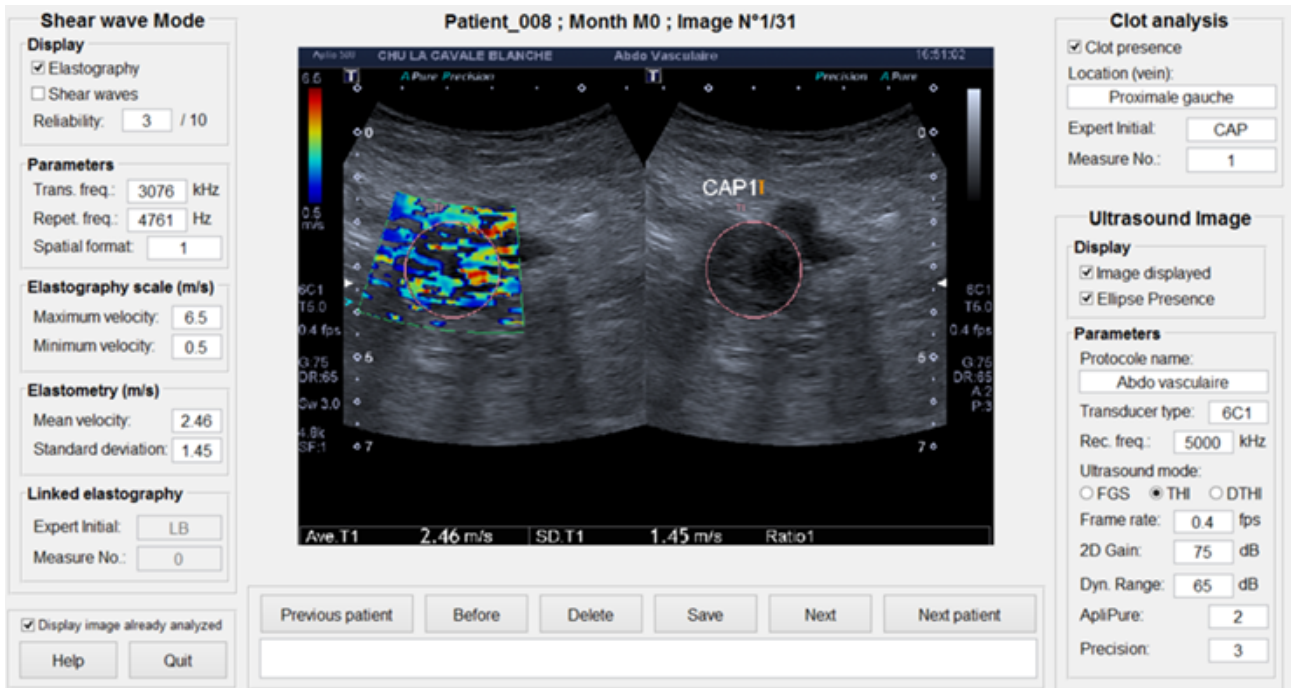
TFD Transformée de Fourier Discrète

Introduction

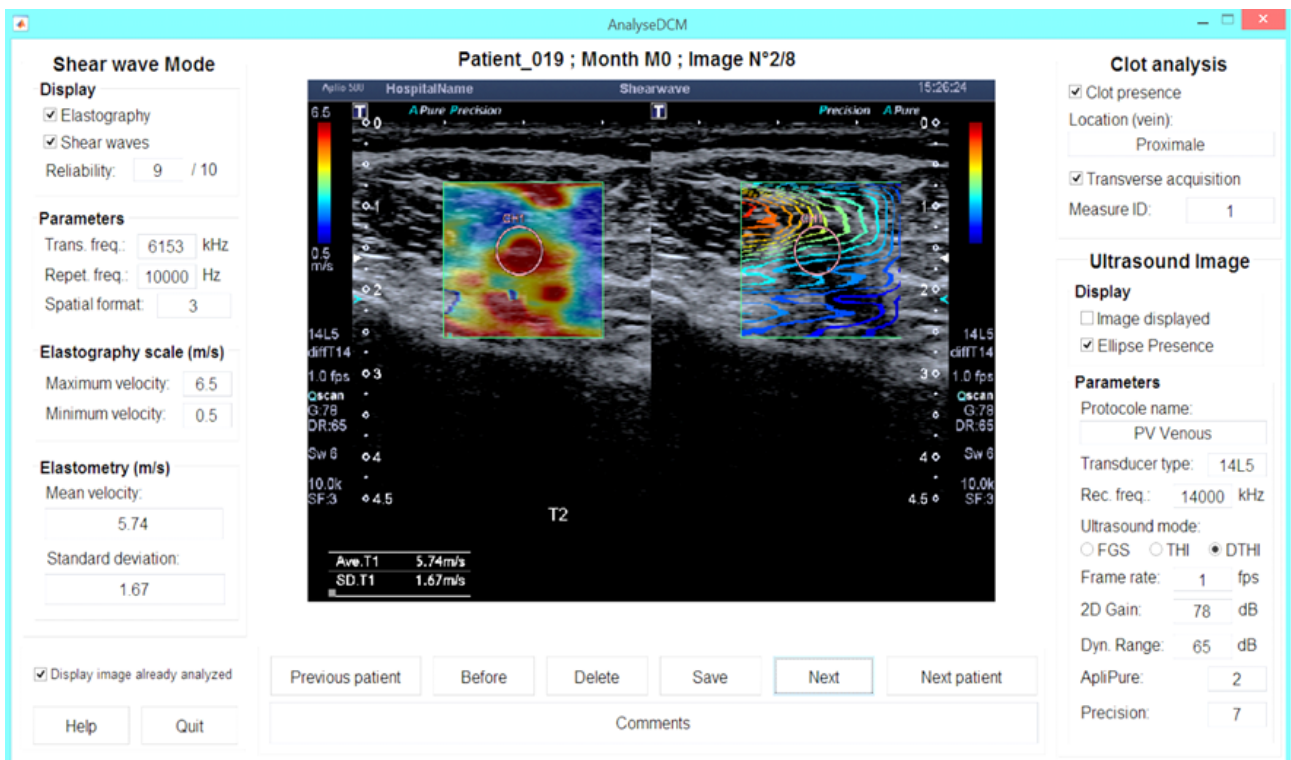
La Base De Données (BDD) a été construite avec le système Toshiba Aplio 500. Nous avons vu à la section 1.3 que le protocole expérimental a évolué avec la mise à jour du système, l'ajout d'acquisitions longitudinales et le retour à la sonde convexe (pour une meilleure stabilité des élastographies). Cet annexe explique en section B.1 l'évolution de l'Interface Homme-Machine (IHM) (présentée en section 3.2.1) qui permet de lire les images, récolter les informations présentes sur l'image et dissocier les deux images de la capture d'écran. La section B.2 propose une méthode par interpolation pour supprimer les annotations des échographies. Les sections B.3 et B.4 détaillent la manière dont les imageries étaient extraites automatiquement à l'intérieur de l'ellipse puis redimensionnées. En effet, à partir du protocole 3, nous avons décidé qu'il serait préférable de ne pas les redimensionner pour préserver la notion d'échelle mais également de prendre en compte le voisinage du thrombus en guise de référence (*cf.* section 3.3.1).

B.1 Lecture et conversion des fichiers DICOM

La mise à jour du système Toshiba Aplio 500 a permis d'améliorer la qualité des élastographies et le nombre d'acquisitions par patient et par opérateur a par conséquent été réduit de dix à trois. Avant la mise à jour, une capture d'écran correspondait à une acquisition : celle-ci comportait donc l'ellipse entourant le thrombus incrustée dans l'image. En diminuant le nombre d'acquisition, nous avons pu augmenter le nombre de captures d'écran sans augmenter la durée de l'examen médical : chaque acquisition comporte alors, en général, une capture sans annotation et une avec annotation. Pour s'adapter au nouveau protocole d'acquisition et récupérer les informations disponibles sur l'image et dans les métadonnées, l'IHM a été modifiée comme on peut le voir sur la figure B.1 :



(a) Version 1



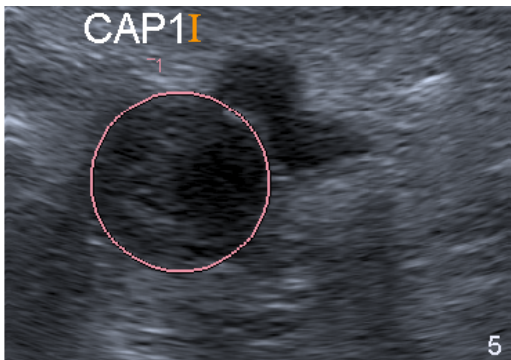
(b) Version 2 : indication de la coupe (transversal ou longitudinal) et possibilité d'entrer plusieurs valeurs d'élastométrie.

FIGURE B.1 – Evolution de l'IHM pour lire les fichiers DICOM (*Digital Imaging and Communications in Medicine*) et récupérer les images, les paramètres d'acquisition et les informations relatives aux patients.

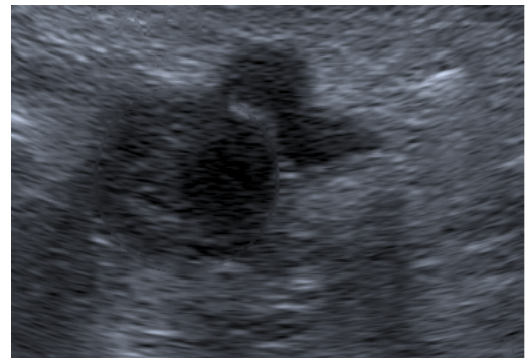
- Une case à cocher a été ajoutée pour indiquer si l’acquisition est transversale ou longitudinale (en haut à droite de l’IHM).
- Le champ « Measure ID » permet de relier les différentes captures d’écran (avec ou sans annotation, avec ou sans carte de propagation) d’une même acquisition. Lors du protocole antérieur, une acquisition comportait le plus souvent une capture d’écran ; une seconde capture avec la carte de propagation était parfois réalisée et pouvait être indiquée dans le champ « Linked elastography » de l’ancienne IHM (champ supprimé dans la nouvelle IHM).
- Les champs d’élastométries ont été agrandis afin de pouvoir entrer plusieurs valeurs moyennes et écart-types dans le cas où il y aurait plusieurs ellipses dessinées (en longitudinal notamment).

B.2 Suppression des annotations et de l’ellipse

Avant d’extraire les images qui seront analysées par la suite, l’ellipse et les annotations sont supprimées manuellement des échographies (voir figure B.2) en utilisant leurs couleurs caractéristiques. Pour chaque pixel X appartenant à un objet à supprimer, la valeur des pixels se trouvant dans une sphère centrée sur X est récupérée. Ensuite, la valeur moyenne des pixels qui appartiennent à la sphère et qui ne font pas partie d’un objet à supprimer est attribuée à X . Toutefois, cette méthode ne fait pas totalement disparaître l’ellipse et l’annotation dont on peut deviner la présence en observant l’image après traitement (voir figure B.2).



(a) Image avec les annotations et l’ellipse



(b) Image après traitement

FIGURE B.2 – Suppression des annotations et de l’ellipse en prenant la valeur moyenne des pixels voisins.

B.3 Extraction d’une imagerie à l’intérieur de l’ellipse

Dans cette section, nous supposons que les différentes images (échographies et élastographies) sont extraites et prétraitées. Ces images contiennent donc le thrombus mais aussi l’artère et les tissus avoisinants. En section 3.3, nous avons vu que plusieurs imageries sont extraites d’une même acquisition et que celles-ci peuvent ne pas être constituées exclusivement de thrombus. Toutefois, dans un premier temps, pour pouvoir caractériser uniquement la structure du thrombus, nous avons utilisé l’ellipse dessinée autour du thrombus pour extraire une imagerie à l’intérieur de ladite ellipse. Cette section présente les trois étapes successives implémentées pour extraire les imageries à l’intérieur des ellipses dessinées par les experts médicaux : la segmentation de l’ellipse, l’extraction de l’imagerie et l’uniformisation de la taille des imageries.

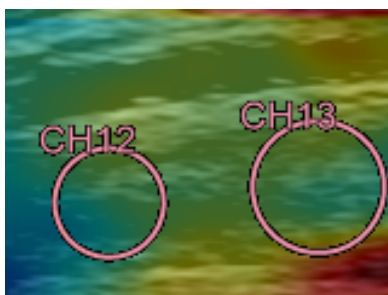
B.3.1 Segmentation des ellipses

Pour extraire une imagerie à l'intérieur de l'ellipse, la première étape consiste à déterminer la position et les paramètres caractéristiques de ladite ellipse. Sur l'image RGB (*Red Green Blue*) de l'échographie ou de l'élastographie, l'ellipse possède des couleurs spécifiques. En fixant à 1 tous les pixels de l'image qui ont ces couleurs et à 0 les autres pixels, on obtient une image binarisée où l'on peut voir le contour de l'ellipse et quelques pixels correspondant principalement aux autres annotations (*cf.* figure B.3.b). Le remplissage du contour de l'ellipse (*cf.* figure B.3.d) suivi d'une érosion (*cf.* figure B.3.e) permet, dans la majorité des cas d'obtenir un masque de l'ellipse (ou des ellipses s'il y en a plusieurs). L'ellipse est donc segmentée et, à partir de fonctions existantes, ses caractéristiques sont calculées : coordonnées du centre, petit axe, grand axe et orientation. Ces paramètres seront utilisés pour extraire, de ces échographies (ou élastographies), les portions d'images qui seront analysées par la suite.

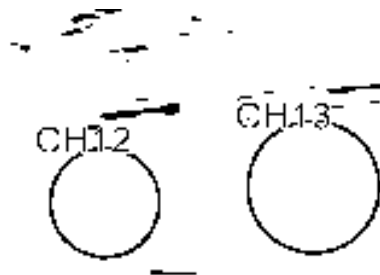
Toutefois, plusieurs cas de figure peuvent se présenter et nécessiter une étape supplémentaire pour segmenter correctement les ellipses :

B.3.1.1 Cas 1 : contour de l'ellipse non fermée

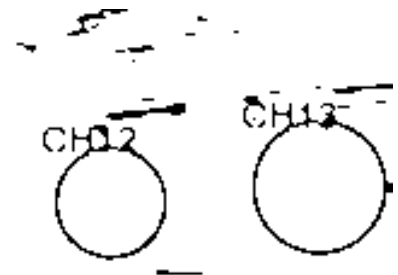
Il arrive parfois que le contour de l'ellipse ne soit pas fermé (présence d'une croix ou de lettres de couleurs différentes sur l'ellipse, ellipse tronquée par le rognage) et que, par conséquent, la fonction de remplissage échoue. Durant la première année, ce cas particulier était traité de façon automatique en ajoutant une étape de fermeture de contour par un élément structurant (croix puis ligne orientée). Avec la présence de plusieurs ellipses, cette méthode automatique peut ne pas détecter la présence d'une des ellipses ou engendrer la fusion de deux ellipses. Une étape de contrôle manuel de la segmentation des ellipses a donc été ajoutée. Elle permet, en cas de mauvaise segmentation, d'indiquer la meilleure stratégie pour fermer le contour : ellipse rognée, croix, ligne orientée. Une fois le contour fermé, les étapes décrites précédemment s'appliquent : remplissage, érosion et récupération des caractéristiques des ellipses. La figure B.3 récapitule les différentes étapes menées pour segmenter deux ellipses dont le contour ne s'est pas fermé correctement. La fermeture avec un élément de type croix a permis de corriger ce problème.



(a) Elastographie avec deux ellipses.



(b) Elastographie binarisée.



(c) Fermeture du contour.



(d) Remplissage des ellipses.



(e) Erosion et étiquetage (chaque ellipse sa couleur).



(f) Produit de l'élastographie et du masque.

FIGURE B.3 – Segmentation des ellipses : cas particulier où le contour n'est pas fermé.

B.3.1.2 Cas 2 : ellipse d'aire insuffisante

L'objectif est d'extraire une imagerie la plus grande possible et de la redimensionner pour quelle est une taille en puissance de 2 (16×16 , 32×32 , 64×64 , 128×128). Dans le cas idéal où l'ellipse est un cercle, l'aire de l'ellipse vaut donc $A = \pi d^2/4$ où d est le diamètre du cercle mais aussi la diagonale du carré extraite. Or, d'après Pythagore, la diagonale d d'un carré de côté c vaut $d = c\sqrt{2}$. Pour extraire une imagerie de taille $c^2 = 32 \times 32$ pixels, il faut donc au moins que l'aire de l'ellipse soit égale à $A = \pi d^2/4 = \pi c^2/2 \approx 1600$ pixels. Afin de limiter les artéfacts liés au redimensionnement de l'imagerie, une étape supplémentaire a été rajoutée dans le cas où l'aire de l'ellipse est inférieure à 1600 pixels. L'IHM qui réalise cette étape est illustrée en figure B.4 : elle permet d'agrandir l'ellipse (seulement si le contenu de l'ellipse reste bien du thrombus) en dessinant une nouvelle région d'intérêt (un polygone), ce qui permet ensuite d'extraire une imagerie plus grande.

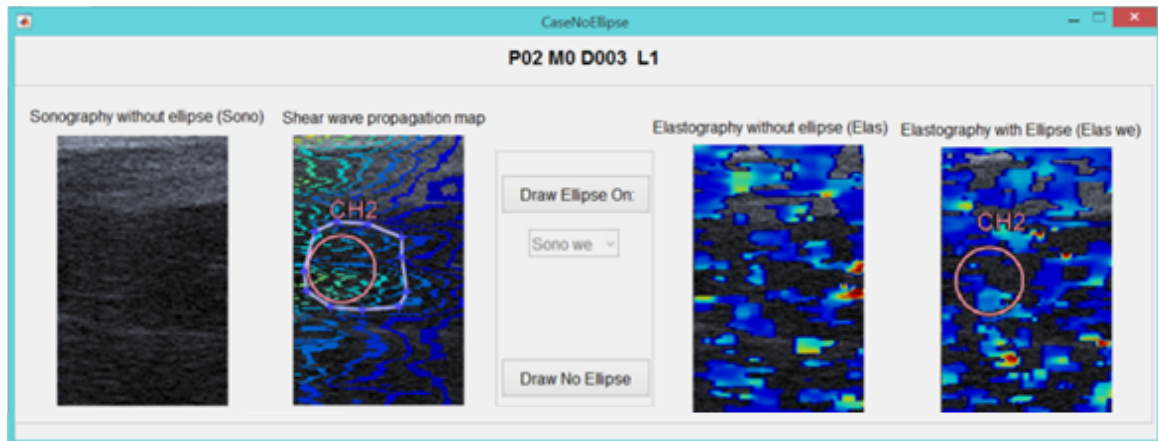


FIGURE B.4 – Cas 3 : agrandissement de l'ellipse.

B.3.1.3 Cas 3 : absence d'ellipse

Dans le cas où l'acquisition ne possède pas d'image avec une ellipse, il est possible de dessiner un polygone autour du thrombus. Ce cas particulier apparaît principalement lorsque l'acquisition en question est une échographie seule (l'élastométrie impose de dessiner une ellipse) comme on peut le voir en figure B.5. La présence du thrombus et sa position peuvent être obtenues en utilisant les élastographies qui suivent cette acquisition.



FIGURE B.5 – Cas 4 : dessin de l'ellipse.

B.3.1.4 Cas 4 : plusieurs ellipses

Enfin, en mode longitudinal, plusieurs ellipses peuvent être dessinées à quelques centimètres d'intervalle afin de mesurer la dureté du thrombus au niveau de sa tête (mesure nommée $L1$) mais aussi un peu plus bas ($L2$ et $L3$). Après avoir segmenté les ellipses, une nouvelle interface permet de définir les ellipses à conserver et leur localisation. Certaines ellipses peuvent être tronquées par le rognage (ou trop petites) et il est donc préférable de ne pas les conserver (ou de les agrandir avant). Cette interface est affichée dans la figure B.6.

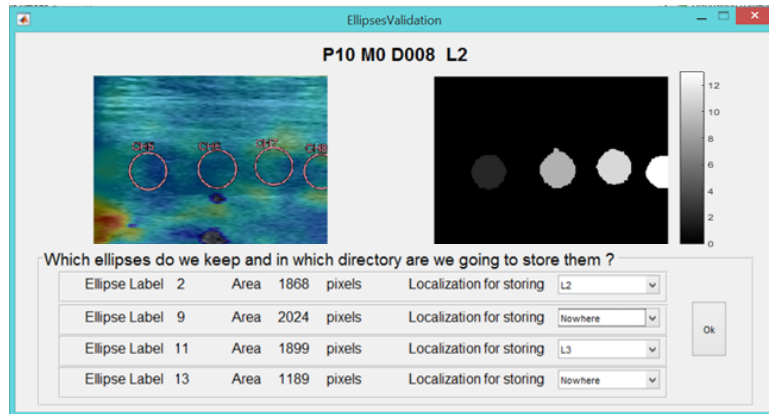
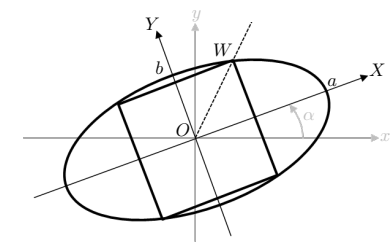


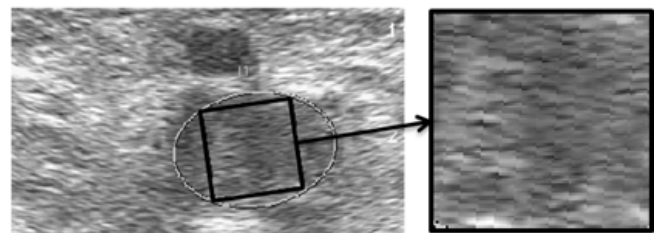
FIGURE B.6 – Cas 4 : dessin de l'ellipse.

B.3.2 Extraction de l'imagette

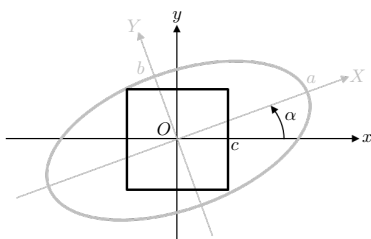
Ici, nous considérons uniquement le thrombus c'est-à-dire que nous allons extraire une image carrée à l'intérieur de l'ellipse. Pour que sa taille soit la plus grande possible, nous avons d'abord appliqué une rotation à l'image de telle sorte que le grand axe et le petit axe de l'ellipse correspondent aux repères cartésiens de l'image après rotation puis nous avons extrait l'image inscrite dans l'ellipse. Toutefois, la rotation de l'image engendre des artefacts sur l'imagette extraite comme nous pouvons le voir en figure B.7.



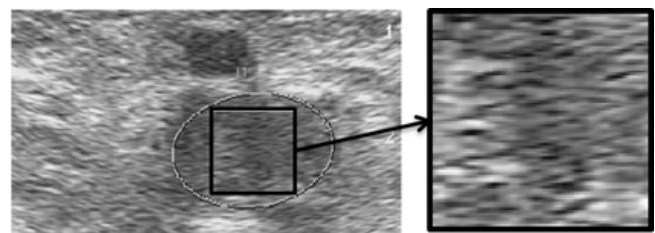
(a) Carré le plus grand possible avec une rotation d'angle α de l'image



(b) Extraction de l'imagette la plus grande possible : artefacts dus à la rotation



(c) Carré le plus grand possible sans rotation de l'image



(d) Extraction de l'imagette la plus grande possible sans rotation : pas d'artéfact mais image légèrement plus petite

FIGURE B.7 – Extraction d'une image carrée la plus grande possible à l'intérieur d'une ellipse

Pour pallier ces artefacts, nous allons extraire une imagette sans appliquer de rotation. Celle-ci sera

donc, sauf cas particulier, plus petite que l'image extraite avec la méthode exposée dans le paragraphe précédent (voir figure B.7). La taille des côtés de l'image carrée vérifiera le système :

$$\frac{X^2}{a^2} + \frac{Y^2}{b^2} = 1 \quad (\text{B.1})$$

où $X = c(\cos \alpha - \sin \alpha)$ and $Y = c(\cos \alpha + \sin \alpha)$. Les images de thrombus extraites sont bien carrées mais ont des tailles différentes suivant la taille de l'ellipse et donc du thrombus.

B.4 Uniformisation de la taille des imagerie

Pour analyser la structure du thrombus, nous allons utiliser un algorithme, nommé *scattering operator*, basé sur la théorie des ondelettes [128]. L'échelle est un paramètre clé des ondelettes et, dans cet algorithme, celle-ci évolue de façon dyadique (en puissance de 2). Pour simplifier les calculs et éviter les approximations, il est préférable d'appliquer l'algorithme à des images dont les dimensions sont des puissances de 2. Les images de thrombus ayant des tailles qui varient de 14×14 à 100×100 pixels pour une taille moyenne d'environ 70×70 pixels, trois tailles cibles seront considérées : 32×32 , 64×64 et 128×128 pixels. Pour redimensionner les imagerie, nous proposons quatre méthodes de redimensionnement :

- **Transformée de Fourier Discrète (TFD)** : L'image est d'abord transformée dans le domaine de Fourier. Si l'image est trop petite nous allons rajouter des zéros sur les bords de la transformée afin de ne pas rajouter d'information dans l'image (*zero padding*). De même, si l'image est trop grande, les pixels des bords, qui contiennent peu d'information, sont supprimés afin de minimiser la perte d'information. On applique ensuite la transformée inverse pour obtenir nos images redimensionnées (cf. section 2.3 et figure 2.4).

- **Transformée en Cosinus Discrète (TCD)** : Le procédé est similaire au précédent mais la TCD vient remplacer la TFD. Dans le domaine des cosinus, l'énergie étant principalement concentrée sur le coin en bas à gauche, les pixels sont ajoutés (ou supprimés) sur les bords droits et hauts de la transformée (cf. section 2.3 et figure 2.4).

- **L'interpolation bilinéaire [54]** : Cette méthode peut être considérée comme une moyenne pondérée de quatre pixels voisins. Elle est assez simple à mettre en œuvre mais elle a l'inconvénient de dégrader les hautes fréquences.

- **L'interpolation bicubique [54]** : Cette méthode suppose que les valeurs entre les pixels voisins varient suivant une cubique (approximation au troisième degré du sinus cardinal). Elle permet d'éviter l'effet passe-bas de l'interpolation bilinéaire.

Pour éviter de changer artificiellement la taille des thrombi en redimensionnement les imagerie, une autre option envisagée est de tout simplement découper les imagerie. Si l'imagerie extraite est de dimension inférieure à la taille cible, alors l'imagerie n'est pas conservée. Si l'imagerie est au moins deux fois plus grande en largeur et en longueur, alors l'imagerie est subdivisée en quatre imagerie. Une imagerie légèrement plus grande que la taille cible est simplement découpée.

Annexe C

Exploration multirésolution des signaux

Introduction

Cette annexe s'intéresse au problème de la représentation des signaux en s'appuyant sur [21, 105, 128]. Les transformations utilisées ont pour objectif de modifier la description d'un signal en le décomposant en « éléments d'information » qui seront, en général, plus simples à interpréter. De nombreuses transformées ont vu le jour ces dernières décennies et cette annexe répertorie les plus usuelles. La section C.1 est consacrée à l'analyse des signaux monodimensionnels et la section C.2 présente l'extension des approches multirésolutions aux signaux bidimensionnels.

C.1 Cas des signaux monodimensionnels

C.1.1 Analyse de Fourier

Soit $s(t)$ un signal à temps continu supposé d'énergie finie :

$$\|s\|^2 = \int_{-\infty}^{+\infty} |s(t)|^2 dt < \infty \quad (\text{C.1})$$

Un outil traditionnel d'analyse de ce signal est sa transformée de Fourier [128] et est définie par :

$$S(f) = \int_{-\infty}^{+\infty} s(t)e^{-i2\pi ft} dt = \langle s(t), e^{i2\pi ft} \rangle \quad (\text{C.2})$$

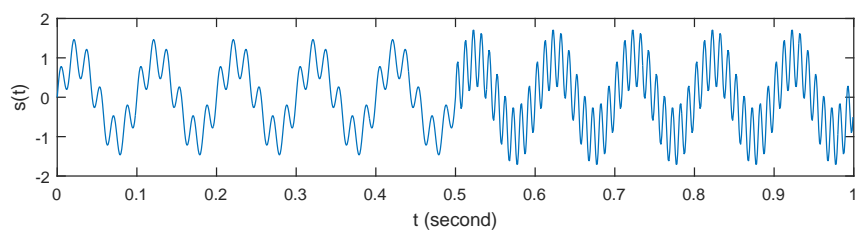
où $\langle a, b \rangle$ désigne le produit scalaire de a et de b .

Cette transformation permet de faire ressortir le contenu fréquentiel du signal mais masque l'évolution temporelle du signal (impulsions, sauts, changements de fréquence ...). Pour pallier cette limitation, il faut construire des représentations faisant intervenir conjointement le temps et la fréquence.

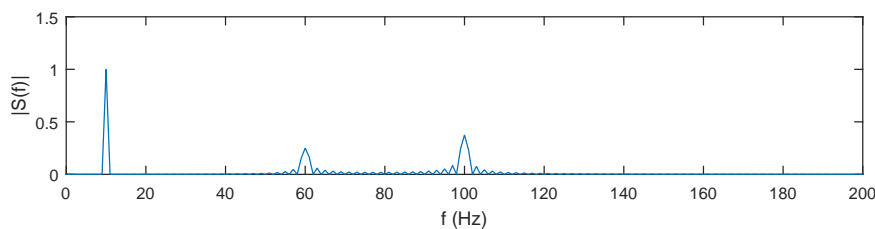
La figure C.1.b montre la transformée de Fourier d'un signal (échantillonné sur 10 000 points avec une fréquence d'échantillonnage de 1 kHz) composé de trois sinusoïdes :

- sinusoïde de fréquence 10 Hz, d'amplitude 1 et durant 1 seconde ;
- sinusoïde de fréquence 60 Hz, d'amplitude 1/2 et durant la première demi-seconde ;
- sinusoïde de fréquence 100 Hz, d'amplitude 3/4 et durant la deuxième demi-seconde.

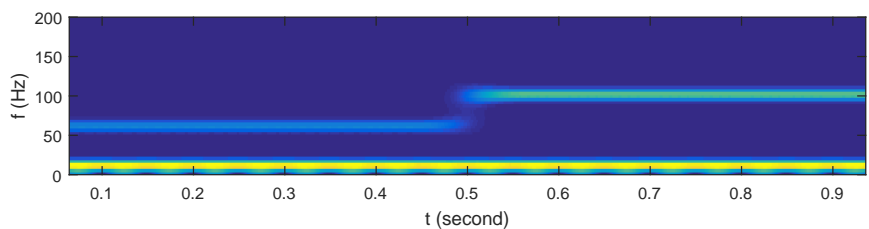
Ce signal, défini en section D.2.1.1 par l'équation (D.68), est illustré en figure C.1.a. La transformée de Fourier permet de retrouver la fréquence des trois sinusoïdes mais pas leur position dans le temps.



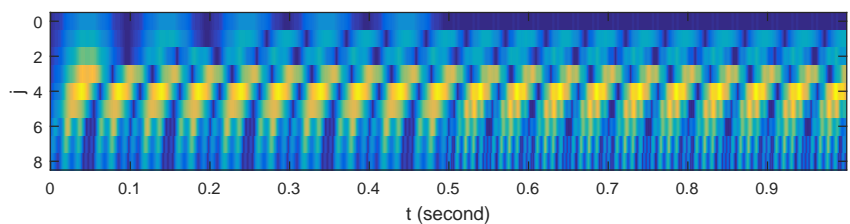
(a) Représentation temporelle



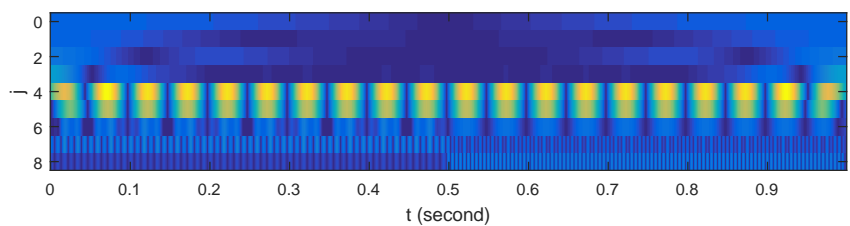
(b) Représentation fréquentielle (Spectre obtenu avec la Transformée de Fourier)



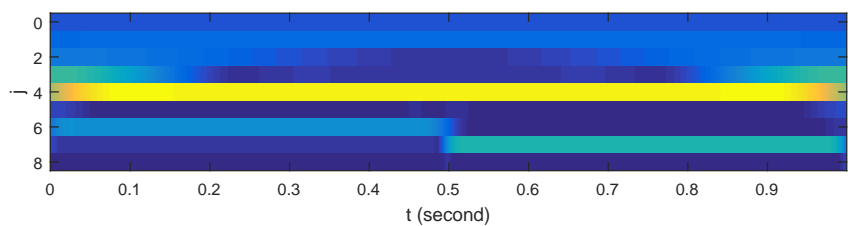
(c) Représentation temps-fréquence (Spectrogramme)



(d) Représentation temps-échelle avec les ondelettes de Haar



(e) Représentation temps-échelle avec les ondelettes du chapeau mexicain



(f) Représentation temps-échelle avec les ondelettes de Morlet

FIGURE C.1 – Analyse d'un signal composée de trois sinusoïdes respectivement de fréquences 10, 60 et 100 Hz , d'amplitudes 1, 1/2 et 3/4 et de durées 1, 1/2 et 1/2 seconde.

C.1.2 Analyse temps-fréquence

Pour analyser un signal non stationnaire, il est d'usage d'établir une représentation temps-fréquence de ce signal. Une telle opération peut s'apparenter à la transcription d'une partition de musique qui indique au musicien les notes (donc l'information fréquentielle) qu'il doit produire à un moment donné.

On parle ici de transformation de Fourier à court terme : on va restreindre l'analyse de Fourier sur une portion du signal délimitée par une fenêtre $g(t)$ localisée autour de l'instant $t = 0$ et on va ensuite faire glisser cette fenêtre le long de l'axe temporel. La transformée de Fourier à court terme [128] s'exprime de la manière suivante :

$$S_c(\nu, \tau) = \int_{-\infty}^{+\infty} s(t)g^*(t - \tau)e^{-i2\pi\nu t} dt = \langle s(t), g(t - \tau)e^{i2\pi\nu t} \rangle \quad (\text{C.3})$$

La quantité $|S_c(\nu, \tau)|^2$ est appelé spectrogramme de $s(t)$ et fournit une indication sur la quantité d'énergie présente dans le signal autour de la fréquence ν et de l'instant τ . Le spectrogramme des trois sinusoïdes en figure C.1.c montre qu'il est possible d'identifier les fréquences et les positions des sinusoïdes dans le temps. On devine également que l'amplitude de la sinusoïde à 10 Hz est supérieure aux deux autres sinusoïdes.

Lors de l'analyse d'un signal, il est important de connaître les localisations en temps et en fréquence des fonctions sur lesquelles on le décompose. D'après le principe d'incertitude de Gabor-Heisenberg [128], il est impossible d'avoir une localisation parfaite à la fois en fréquence et en temps. Toutefois, on peut montrer que les fenêtres conduisant à la meilleure localisation temps-fréquence possible sont de forme gaussienne. On parle alors d'analyse de Gabor. Un des avantages de la transformée de Fourier à court terme est qu'elle fournit une représentation complète du signal ; ce dernier peut être reconstruit à l'identique à partir de sa représentation temps-fréquence.

C.1.3 Analyse temps-échelle

La transformée en ondelette constitue [128] (ou analyse temps-échelle) une alternative intéressante à l'analyse temps-fréquence. Elle est définie par la formule suivante :

$$W_s(\alpha, b) = |\alpha|^{1/2} \int_{-\infty}^{+\infty} s(t)\psi^*(\alpha(t - b))dt = \langle s(t), |\alpha|^{1/2}\psi(\alpha(t - b)) \rangle \quad (\text{C.4})$$

où $\psi(t)$ est une fonction d'énergie finie et $|W_s(\alpha, b)|^2$ est nommée scalogramme de $s(t)$.

$\{|\alpha|^{1/2}\psi(\alpha(t - b)), (\alpha, b) \in \mathbb{R}^* \times \mathbb{R}\}$ est une famille de fonctions formée de contractées/dilatées (suivant que $|\alpha| < 1$ ou $|\alpha| > 1$) et translatées de $\psi(t)$. On dit qu'il s'agit d'une famille d'ondelettes dont ψ est l'ondelette mère.

Si on suppose que $\psi(t)$ est associée au pavé $[-\Delta t/2, \Delta t/2] \times [\bar{f} - \frac{\Delta f}{2}, \bar{f} + \frac{\Delta f}{2}]$ dans le demi-plan temps-fréquence correspondant aux fréquences positives, \bar{f} étant la fréquence moyenne sur le demi-axe des fréquences positives. L'ondelette fille $|\alpha|^{1/2}\psi(\alpha(t - b))$ est localisée temporellement autour de l'instant b , nommé paramètre de localisation.

On parle de représentation temps-échelle, car, lorsque α varie, l'ondelette change d'échelle pour pouvoir couvrir différentes bandes de fréquences. Par exemple, si $\alpha < 1$, la dilatée $\psi(\alpha t)$ a un support plus étroit en fréquence, $[(\alpha\bar{f}) - \alpha\frac{\Delta f}{2}, (\alpha\bar{f}) + \alpha\frac{\Delta f}{2}]$, mais plus étalé en temps, $[-\frac{\Delta t}{2\alpha}, \frac{\Delta t}{2\alpha}]$. Le paramètre $1/|\alpha|$ est nommé facteur d'échelle : il correspond à des échelles fines (hautes fréquences) lorsqu'il est élevé et à des échelles grossières (basses fréquences) lorsqu'il est faible. En pratique, les échelles des transformées en ondelettes sont échantillonnées suivant une suite géométrique de raison 2. On parle alors de transformées en ondelettes dyadiques et le facteur d'échelle $1/|\alpha|$ est défini en fonction de l'entier naturel j par $\alpha = 2^{-j}$.

Parmi les ondelettes usuelles (cf. figure C.2), on retrouve :

- **Haar :**

$$\psi(t) = \begin{cases} 1 & \text{pour } 0 < t < 1/2 \\ -1 & \text{pour } 1/2 < t < 1 \\ 0 & \text{sinon} \end{cases} \quad (\text{C.5})$$

- **Dérivée n -ième de la Gaussienne :**

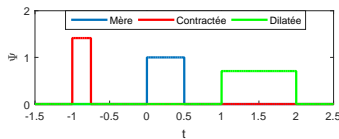
$$\psi(t) = \frac{d^n}{dt^n} \exp(-\pi t^2), \quad n \geq 1 \quad (\text{C.6})$$

- **Chapeau mexicain** (proportionnelle à la dérivée seconde de la fonction de Gauss) :

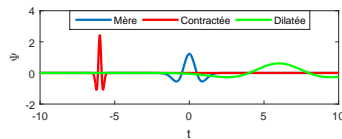
$$\psi(t) = \frac{2}{\pi^{1/4} \sqrt{3\sigma}} \left(1 - \frac{t^2}{\sigma^2}\right) \exp\left(-\frac{t^2}{2\sigma^2}\right), \quad \sigma > 0 \quad (\text{C.7})$$

- **Morlet :**

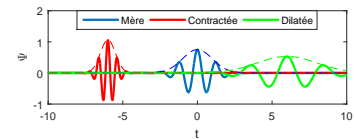
$$\psi(t) = \frac{1}{\pi^{1/4} \sqrt{\sigma}} \exp\left(-\frac{t^2}{2\sigma^2} + i2\pi \bar{f}t\right), \quad \sigma > 0 \quad (\text{C.8})$$



(a) Ondelettes de Haar



(b) Ondelettes du chapeau mexicain



(c) Ondelettes de Morlet (partie réelle en trait plein et module en pointillés)

FIGURE C.2 – Représentation de l'ondelette mère (au centre), d'une ondelette fille contractée (à gauche) et d'une dilatée (à droite) pour trois familles d'ondelette : Haar, chapeau mexicain et Morlet.

Le principal avantage de l'analyse temps-échelle, par rapport à l'analyse de Fourier à court terme par exemple, est qu'elle se révèle souvent mieux adaptée aux caractéristiques des signaux naturels et à notre façon de les percevoir. Sur la figure C.1, on peut voir une application de ces trois ondelettes à un signal composé de trois sinusoïdes. L'ondelette de Morlet permet de retrouver facilement l'échelle ($j = 4$ pour la sinusoïde de fréquence 10 Hz , $j = 6$ pour la sinusoïde de fréquence 60 Hz et $j = 7$ pour la sinusoïde de fréquence 100 Hz). Plus l'échelle est élevée, plus la fréquence correspondante est élevée également. La sinusoïde de 10 Hz se retrouve aussi avec les ondelettes du chapeau mexicain et de Haar mais les deux autres sinusoïdes sont plus compliquées à identifier avec celle de Haar.

De plus, comme en temps-fréquence, il est possible de reconstruire le signal à partir de $W_s(\alpha, b)$:

$$s(t) = \frac{1}{K_\Psi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} W_s(\alpha, b) |\alpha|^{1/2} \psi(\alpha(t-b)) d\alpha db \quad (\text{C.9})$$

où $K_\Psi = \int_{-\infty}^{+\infty} \frac{|\Psi(f)|^2}{|f|} df$ et $\Psi(f)$ la transformée de Fourier de l'ondelette mère. Cette formule d'inversion nécessite que la constante K_ψ soit finie, ce qui est appelée condition d'admissibilité et ce qui nécessite que l'ondelette mère soit de moyenne nulle :

$$\Psi(0) = \int_{-\infty}^{+\infty} \psi(t) dt = 0 \quad (\text{C.10})$$

Les ondelettes de Haar (C.5) et du chapeau mexicain (C.7) sont rigoureusement admissibles alors que l'ondelette de Morlet C.8 l'est approximativement si l'on impose une valeur suffisamment faible à $\Psi(0)$ (typiquement $2\pi\sigma\bar{f} \approx 5$).

Toutefois, l'expression de $s(t)$ en fonction de $W_s(\alpha, b)$ fait intervenir les ondelettes correspondant à toutes les localisations temporelles et à tous les facteurs d'échelle, ce qui la rend difficile à réaliser. C'est pourquoi la section suivante va présenter une analyse avec un nombre fini d'échelles appelée aussi analyse multirésolution.

C.1.4 Analyse multirésolution

L'Analyse MultiRésolution (AMR) [128] consiste à décomposer le signal sur un nombre fini d'ondelettes. Cela est notamment possible en faisant varier le facteur d'échelle de façon dyadique c'est-à-dire en choisissant $\alpha = 2^{-j}$ et $b = k/\alpha$, $j, k \in \mathbb{Z}^1$.

Dans ce cas-là, l'Analyse MultiRésolution (AMR) est définie comme la suite de sous-espaces vectoriels fermés $(V_j)_{j \in \mathbb{Z}}$ de l'espace $L^2(\mathbb{R})$ des signaux d'énergie finie et vérifiant les quatre propriétés suivantes :

- $(V_j)_{j \in \mathbb{Z}}$ constitue une suite d'espaces emboîtés :

$$\forall j \in \mathbb{Z}, V_{j+1} \subset V_j \quad (\text{C.11})$$

La projection sur V_{j+1} constitue nécessairement une moins bonne approximation du signal que la projection sur V_j .

- Un signal $s(t)$ appartient à V_j si et seulement si sa dilatée $s(t/2)$ appartient à V_{j+1} :

$$\forall j \in \mathbb{Z}, s(t) \in V_j \Leftrightarrow s(t/2) \in V_{j+1} \quad (\text{C.12})$$

En d'autres termes, on passe de V_j à V_{j+1} en multipliant par 2 le facteur d'échelle. C'est pourquoi j est nommé niveau de résolution.

- Conditions aux limites :

$$\lim_{j \rightarrow \infty} V_j = \bigcap_{j \rightarrow -\infty}^{j \rightarrow +\infty} V_j = \{\emptyset\} \quad (\text{C.13})$$

$$\lim_{j \rightarrow -\infty} V_j = \bigcup_{j \rightarrow -\infty}^{j \rightarrow +\infty} V_j = L^2(\mathbb{R}) \left(\text{i.e. que } \lim_{j \rightarrow -\infty} V_j \text{ est dense dans } L^2(\mathbb{R}) \right) \quad (\text{C.14})$$

On perd donc toutes les informations (\emptyset représente l'ensemble vide) lorsque j tend vers ∞ alors que, lorsque j tend vers $-\infty$, l'approximation est asymptotiquement égale au signal d'origine.

- Il existe une fonction $\varphi(t) \in L^2(\mathbb{R})$ qu'on appelle fonction d'échelle ou ondelette père telle que $\{\varphi(t-k), k \in \mathbb{Z}\}$ forme une base orthonormale de V_0 . Ceci implique que :

$$\forall k \in \mathbb{Z}, \varphi(t) \in V_j \Rightarrow \varphi(t - k2^j) \in V_j \quad (\text{C.15})$$

Les conditions 3 et 4 se traduisent par le fait que $\{2^{-j/2}\varphi(t2^{-j} - k), k \in \mathbb{Z}\}$ forment une base orthogonale de V_j . On parle alors de base de Riesz.

L'AMR consiste à réaliser des projections orthogonales successives du signal sur les espaces V_j ce qui conduit à des approximations de plus en plus grossières du signal au fur et à mesure que le niveau de résolution j augmente. On peut faire l'analogie avec un cameraman qui effectuerait un « zoom » : il part d'une vision globale des choses puis il va se focaliser sur des zones d'intérêt en faisant apparaître plus de détails par rapport à la première image.

La différence entre deux approximations consécutives représente l'information de « détail » qui est perdue au passage d'une échelle à l'autre ; cette information est contenue dans un sous-espace W_j orthogonal à V_j tel que $V_{j-1} = V_j \oplus W_j$. Il existe donc une ondelette mère $\psi(t) \in L^2(\mathbb{R})$ telle que $\{2^{-j/2}\psi(t2^{-j} - k), k \in \mathbb{Z}\}$ forme une base orthonormale de W_j . Par union de toutes ses familles, quand j varie, on peut montrer que $\{2^{-j/2}\psi(t2^{-j} - k), (j, k) \in \mathbb{Z}^2\}$ constitue une base orthonormale d'ondelettes de $L^2(\mathbb{R})$.

1. Ensemble des entiers relatifs

La décomposition en ondelettes orthogonales d'un signal $s(t)$ peut être effectuée de manière très efficace avec l'algorithme de Mallat en déterminant pour chaque niveau de résolution $j \in \mathbb{Z}$ les coefficients d'approximations $(a_j[k])_{k \in \mathbb{Z}}$ et de détails $(c_j[k])_{k \in \mathbb{Z}}$ définis par :

$$\forall k \in \mathbb{Z}, a_j[k] = \langle s(t), 2^{-j/2} \varphi(t/2^j - k) \rangle \quad (\text{C.16})$$

$$c_j[k] = \langle s(t), 2^{-j/2} \psi(t/2^j - k) \rangle \quad (\text{C.17})$$

Avant de s'intéresser à l'analogie entre la théorie des ondelettes et celle des bancs de filtres (cf. section C.1.5) dont découle l'algorithme de Mallat (cf. section C.1.6), on va d'abord voir un exemple de décomposition d'une fonction en escalier sur la base de Haar sur $L^2(0, 1)$:

$$\forall t \in [0, 1], \varphi(t) = 1 \text{ et } \psi(t) = \begin{cases} 1 & \text{pour } 0 \leq t < 1/2 \\ -1 & \text{pour } 1/2 \leq t < 1 \end{cases} \quad (\text{C.18})$$

Pour $j \geq 0$ et $0 \leq k \leq 2^j - 1$, la décomposition en ondelettes orthogonales $\psi_{(j,k)}(t) = 2^{j/2} \psi(2^j t - k)$ donne, dans le cas particulier de l'ondelette de Haar, la décomposition suivante :

$$\psi_{(j,k)}(t) = \begin{cases} 2^{j/2} & \text{si } t \in [k 2^{-j}, (k + 1/2)2^{-j}] \\ -2^{j/2} & \text{si } t \in [(k + 1/2) 2^{-j}, (k + 1)2^{-j}] \end{cases} \quad (\text{C.19})$$

La fonction en escalier $[2, 4, 7, 11, 12, 0, 3, 1]$ sur $[0, 1]$, après décomposition en ondelettes de Haar avec trois niveaux de résolutions est équivalente à l'expression $[5, 1, -3, 2, -1, -2, 6, 1]$. La décomposition est visible sur la figure C.3 avec les coefficients d'approximation en bleu et de détails en rouge.

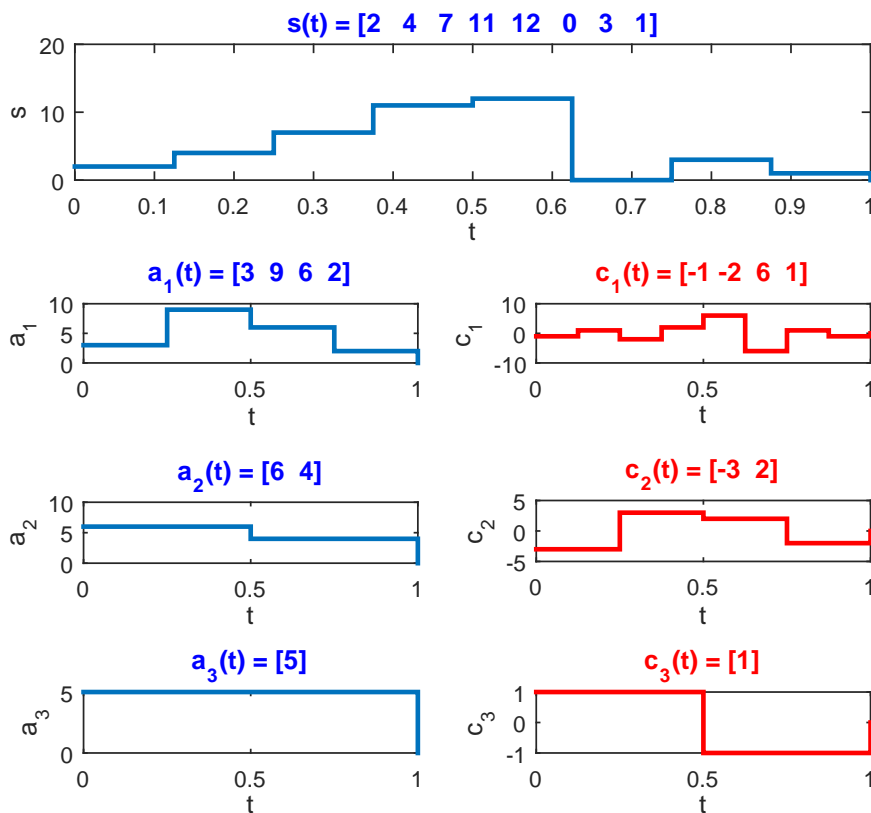


FIGURE C.3 – Décomposition de Haar suivant trois niveaux de résolution j d'une fonction en escalier $s(t)$: coefficients de d'approximation $a_j(t)$ à gauche et de détails $c_j(t)$ à droite.

C.1.5 Ondelettes et bancs de filtres

L'étude des propriétés des ondelettes orthogonales et de l'analyse multirésolution a mis en lumière un lien avec les bancs de filtres. Cette section s'intéresse à la construction de bancs de filtres d'analyse en quadrature miroir [128]. La mise en cascade de tels bancs de filtres conduit à une décomposition du signal sur plusieurs niveaux de résolution (voir figure C.4).

D'après la propriété n° 1 de l'AMR, pour un entier relatif j , l'ensemble V_{j+1} est inclus dans V_j donc on peut écrire toute fonction de V_{j+1} comme la combinaison linéaire des fonctions $\varphi(t-l)$ tel que $l \in \mathbb{Z}$ qui forment une base de V_j . Soit k un entier relatif et $2^{-(j+1)/2} \varphi(t 2^{-(j+1)} - k)$, une fonction de V_{j+1} , alors il existe une suite $(h_0[l])_{l \in \mathbb{Z}}$ telle que :

$$2^{-(j+1)/2} \varphi(t 2^{-(j+1)} - k) = \sum_{l=-\infty}^{+\infty} h_0[l - 2k] 2^{-j/2} \varphi(t 2^{-j} - l) \quad (\text{C.20})$$

L'équation (C.20) est appelée équation à 2 échelles. La suite $(h_0[l])_{l \in \mathbb{Z}}$ peut être vue comme la réponse impulsionnelle d'un filtre dont la réponse fréquentielle serait sa transformée de Fourier :

$$H_0[f] = \sum_{l=-\infty}^{l=+\infty} h_0[l] e^{-i2\pi fl} \quad (\text{C.21})$$

Cette dernière est une fonction périodique (de période 1) et on peut montrer que $H_0[0] = \sqrt{2}$.

De la même façon, on peut raisonner sur les espaces de détails $W_{j+1} \subset V_j$ où les ondelettes $\{2^{-(j+1)/2} \psi(t 2^{-(j+1)} - k), k \in \mathbb{Z}\}$ forment une base de W_{j+1} et peuvent être décomposées sur la base orthonormale $\{2^{-j/2} \varphi(t 2^{-j} - l), l \in \mathbb{Z}\}$ de V_j selon la relation, $\forall k \in \mathbb{Z}$:

$$2^{-(j+1)/2} \psi(t 2^{-(j+1)} - k) = \sum_{l=-\infty}^{+\infty} h_1[l - 2k] 2^{-j/2} \varphi(t 2^{-j} - l) \quad (\text{C.22})$$

La suite $(h_1[l])_{l \in \mathbb{Z}}$ peut aussi s'interpréter comme la réponse impulsionnelle d'un filtre dont on peut montrer que $H_1[0] = 0$. Cette condition couplée à la condition $H_0[0] = \sqrt{2}$ implique que les filtres de réponses impulsionnelles $(h_0[l])_{l \in \mathbb{Z}}$ et $(h_1[l])_{l \in \mathbb{Z}}$ doivent être respectivement passe-bas et passe-haut [128].

C.1.6 Algorithme de Mallat

L'analogie entre la théorie des ondelettes et celle des bancs de filtres permet d'effectuer une décomposition en ondelettes de manière efficace [22, 128]. Pour ce faire, on détermine, à chaque niveau de résolution j , les projections sur l'espace des approximations $\text{proj}_{V_j}[s(t)]$ et sur celui des informations de détails $\text{proj}_{W_j}[s(t)]$ du signal $s(t)$. Les familles $\{2^{-j/2} \varphi(t 2^{-j} - k), k \in \mathbb{Z}\}$ et $\{2^{-j/2} \psi(t 2^{-j} - k), k \in \mathbb{Z}\}$ étant des bases orthonormales respectives des espaces V_j et W_j , on a :

$$\text{proj}_{V_j}[s(t)] = \sum_{k=-\infty}^{+\infty} a_j[k] 2^{-j/2} \varphi(t 2^{-j} - k) \quad (\text{C.23})$$

$$\text{proj}_{W_j}[s(t)] = \sum_{k=-\infty}^{+\infty} c_j[k] 2^{-j/2} \psi(t 2^{-j} - k) \quad (\text{C.24})$$

avec $a_j[k]$ et $c_j[k]$ tels que définis aux équations (C.16) et (C.17).

Le calcul direct par intégration sur \mathbb{R} de l'ensemble des produits scalaires des coefficients d'approximation $(a_j[k])_{k \in \mathbb{Z}}$ (C.16) et de détails $(c_j[k])_{k \in \mathbb{Z}}$ (C.17) serait fastidieux. Toutefois, les équations à 2 échelles (C.20) et (C.22) montrent des relations récursives qui simplifient grandement les calculs :

$$a_j[k] = \sum_{l=-\infty}^{+\infty} a_{j-1}[l] h_0^*[l - 2k] \quad (\text{C.25})$$

$$c_j[k] = \sum_{l=-\infty}^{+\infty} a_{j-1}[l] h_1^*[l - 2k] \quad (\text{C.26})$$

où $h_0^*[l - 2k]$ désigne le conjugué de $h_0[l - 2k]$.

A partir des coefficients d'approximation du signal à une échelle donnée 2^{j-1} , on peut ainsi déterminer les coefficients d'approximation et de détails à l'échelle deux fois plus grande. A chaque niveau de résolution, on va effectuer un traitement par filtre de réponse impulsionnelle $(h_0^*[k])_{k \in \mathbb{Z}}$ et $(h_1^*[k])_{k \in \mathbb{Z}}$ suivi d'un sous-échantillonnage d'un facteur 2 (symbolisé par $2 \downarrow$ sur la figure C.4). On obtient alors un banc de filtres d'analyse en quadrature miroir schématisé par la figure C.4.

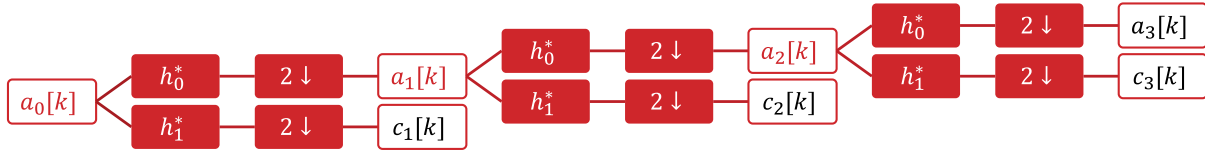


FIGURE C.4 – Structure de décomposition en ondelettes sur 3 niveaux de résolution.

A partir du dernier coefficient d'approximation et des coefficients de détails, il est possible de reconstruire le signal d'origine en utilisant un banc de filtres de synthèse qui réalise les opérations duales de celles effectuées lors de la décomposition : une interpolation de facteur 2 et un filtrage.

C.1.7 Généralisation à M -bandes

Dans les sections précédentes, on a fait varier le facteur d'échelle de façon dyadique. Néanmoins, il est également possible de le faire varier sous la forme M^j avec $M \geq 2$ [29]. Le cas dyadique est le cas particulier pour $M = 2$. On part d'espaces V_j et $(W_j^m)_{m \in \{1, 2, \dots, M-1\}}$ définis à un niveau de résolution j , qui sont liés par la relation :

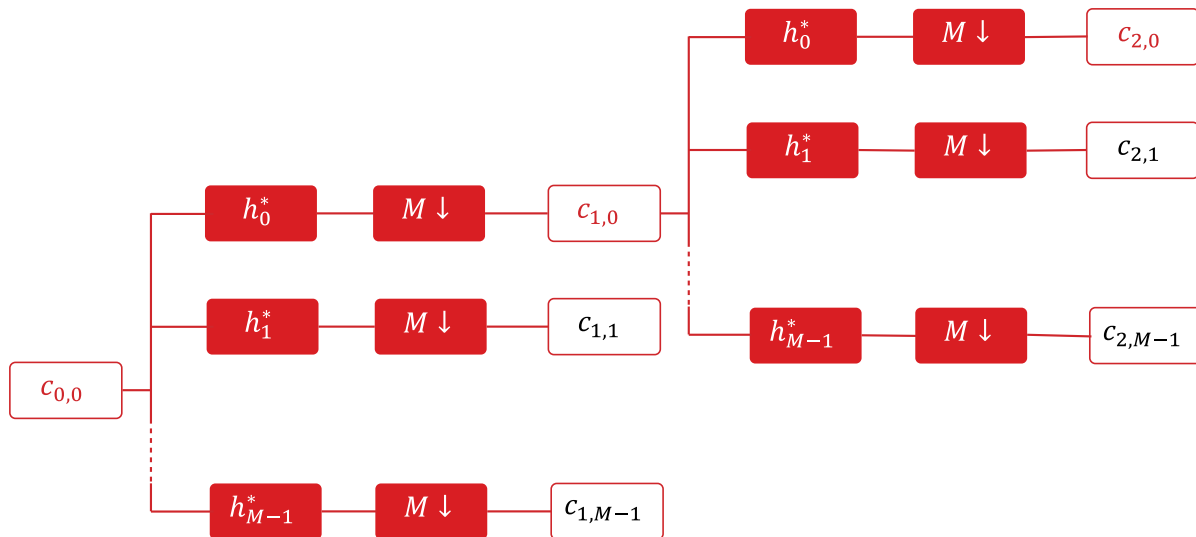
$$V_{j-1} = V_j \oplus \left(\bigoplus_{m=1}^{M-1} W_j^m \right) \quad (\text{C.27})$$

L'équation (C.27) signifie que l'espace V_{j-1} peut être décomposé en sous-espaces orthogonaux V_j et $(W_j^m)_{m \in \{1, 2, \dots, M-1\}}$.

Les familles $\{M^{-j/2} \varphi(tM^{-j} - k), k \in \mathbb{Z}\}$ et $\{M^{-j/2} \psi_m(tM^{-j} - k), k \in \mathbb{Z}\}$ sont respectivement des bases orthonormales de V_j et de W_j^m . Une analyse multirésolution M -bandes de $L^2(\mathbb{R})$ met donc en œuvre une fonction d'échelle $\varphi(t) = \psi_0(t) \in L^2(\mathbb{R})$ et $M - 1$ ondelettes mère $\psi_m \in L^2(\mathbb{R})$ qui sont les solutions de l'équation à deux échelles suivantes, $\forall m \in \{0, 1, \dots, M - 1\}$:

$$\frac{1}{\sqrt{M}} \psi_m \left(\frac{t}{M} - k \right) = \sum_{k=-\infty}^{+\infty} h_m[k] \varphi(t - k) \quad (\text{C.28})$$

La reconstruction à partir du dernier coefficient d'approximation et des coefficients de détails est également possible. La figure C.5 illustre une analyse multirésolution M -bandes sur 2 niveaux de résolutions.


 FIGURE C.5 – Analyse multirésolution M -bandes sur 2 niveaux de résolution.

C.1.8 Paquet d'ondelettes

La décomposition en paquets d'ondelettes [29, 128] a pour objectif de s'adapter au mieux aux caractéristiques temps-fréquence des signaux. Contrairement à la transformée en ondelettes discrète vue précédemment et qui décompose uniquement les sous-bandes passe-bas, elle va également itérer ce processus de décomposition sur les sous-bandes haute fréquence (comme on peut le voir sur la figure C.6).

Dans le cas à 2 bandes, on va définir les fonctions $(\psi_m)_{m \in \mathbb{N}}$ (de transformée de Fourier $(\Psi_m)_{m \in \mathbb{N}}$) telles que :

$$\forall m \in \mathbb{N}, \begin{cases} \sqrt{2} \Psi_{2m}(2\omega) = A_0(\omega) \Psi_m(\omega) \\ \sqrt{2} \Psi_{2m+1}(2\omega) = A_1(\omega) \Psi_m(\omega) \end{cases} \quad (\text{C.29})$$

où A_0 et A_1 sont les réponses fréquentielles des filtres passe-bas du banc de filtres 2-bandes associé. Soit W_j^m le sous espace vectoriel dont la famille $\{2^{-j/2} \psi_m(t 2^{-j} - k), k \in \mathbb{Z}\}$ est une base, alors, au niveau de résolution $j + 1$, on obtiendra les familles $\{2^{-(j+1)/2} \psi_{2m}(t 2^{-(j+1)} - k), k \in \mathbb{Z}\}$ et $\{2^{-(j+1)/2} \psi_{2m+1}(t 2^{-(j+1)} - k), k \in \mathbb{Z}\}$ bases orthonormales respectives des espaces vectoriels W_{j+1}^{2m} et W_{j+1}^{2m+1} vérifiant :

$$W_j^m = W_{j+1}^{2m} \oplus W_{j+1}^{2m+1} \quad (\text{C.30})$$

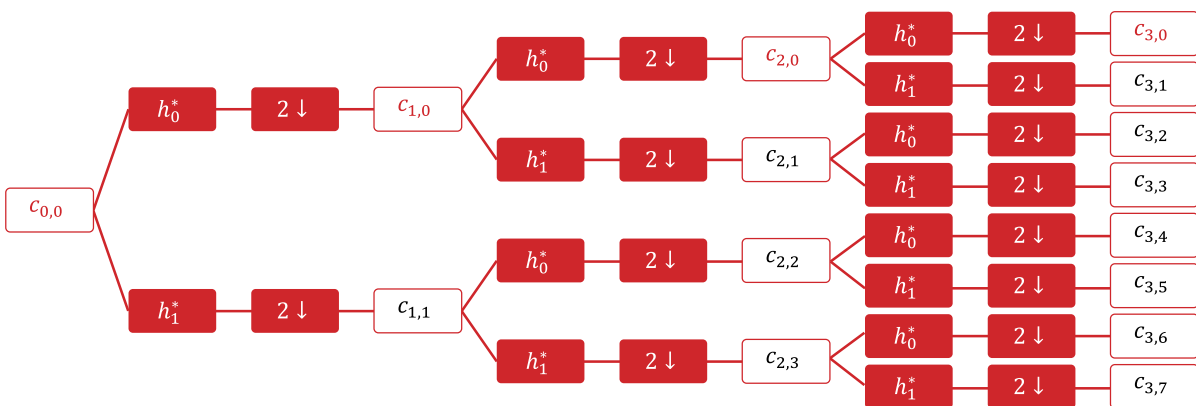


FIGURE C.6 – Structure de décomposition en paquets d'ondelettes sur 3 niveaux de résolution.

De plus, des ondelettes M -bandes peuvent être générées à partir de décompositions en paquets d'ondelettes dyadiques correspondant à une analyse en sous-bandes égales, ce qui signifie que toutes les sous-bandes obtenues seront identiquement recomposées sur les différents niveaux de résolution. De même, la reconstruction de l'image d'origine à partir des coefficients d'ondelettes est possible.

C.2 Cas des signaux bidimensionnels

Contrairement à la section précédente qui considère des transformées s'appliquant à des signaux monodimensionnels, cette section est consacrée à l'usage des ondelettes sur des signaux bidimensionnels : des images. Les ondelettes peuvent être séparables (comme la transformée en ondelettes 2D usuelle) ou non séparables. Le but est toujours d'extraire les informations, les caractéristiques ou les détails essentiels présents dans les données. Sur une image, les coefficients d'ondelettes les plus énergétiques indiquent la position des contours, où l'intensité lumineuse varie brusquement (utilisés en reconnaissance de forme en vision par ordinateur). Le zoom de la transformée en ondelettes permet de localiser et caractériser des singularités, qu'elles soient isolées ou non (multifractale [128]).

C.2.1 Transformée en ondelettes bidimensionnelles séparables

L'analyse multirésolution bidimensionnelle (2D) [18, 128] est une extension directe du cas monodimensionnel (1D) (*cf.* section C.1) puisqu'elle consiste à traiter l'image suivant chaque dimension de l'espace : on procède successivement à une analyse des lignes puis des colonnes de l'image comme on peut le voir avec la structure associée de bancs de filtres séparables représentées sur la figure C.7. Les ondelettes 2D sont les produits tensoriels des ondelettes 1D.

Si $(V_j)_{j \in \mathbb{Z}}$ est une AMR de $L^2(\mathbb{R})$, alors on peut en déduire une AMR séparable de $L^2(\mathbb{R}^2)$ notée $(V_j^2)_{j \in \mathbb{Z}}$. Soit $\varphi(t)$ une fonction d'échelle de V_j^2 , alors une base orthogonale de V_j^2 est donnée par, $\forall (n, m) \in \mathbb{Z}^2$ et $(x, y) \in \mathbb{R}^2$:

$$2^{-j} \varphi(x 2^{-j} - n) \varphi(y 2^{-j} - m) \quad (\text{C.31})$$

Ensuite, on va chercher, comme dans le cas monodimensionnel, une base d'ondelettes dans l'espace vectoriel qui est le complémentaire orthogonal de V_j^2 dans V_{j-1}^2 . Mais, en 2D, on va définir, non plus un, mais trois espaces vectoriels de détails W_j^H , W_j^V et W_j^D de telle sorte que :

$$V_{j-1}^2 = V_j^2 \oplus W_j^H \oplus W_j^V \oplus W_j^D \quad (\text{C.32})$$

Si l'on choisit $\psi(t)$ comme ondelette mère associée à l'AMR $(V_j)_{j \in \mathbb{Z}}$, on va alors pouvoir définir :

- L'ondelette directionnelle $\psi_H(x, y) = \varphi(x)\psi(y)$ dans la direction horizontale (H) ;
- L'ondelette directionnelle $\psi_V(x, y) = \psi(x)\varphi(y)$ dans la direction verticale (V) ;
- L'ondelette directionnelle $\psi_D(x, y) = \psi(x)\psi(y)$ dans la direction diagonale (D).

On va obtenir des bases orthonormales des trois sous-espaces vectoriels de détails W_j^H , W_j^V et W_j^D par translation et dilatations/contractions des fonctions $\psi_H(x, y)$, $\psi_V(x, y)$ et $\psi_D(x, y)$:

- $\{2^{-j} \psi_H(x 2^{-j} - n, y 2^{-j} - m), (n, m) \in \mathbb{Z}^2\}$ base orthonormale de W_j^H ;
- $\{2^{-j} \psi_V(x 2^{-j} - n, y 2^{-j} - m), (n, m) \in \mathbb{Z}^2\}$ base orthonormale de W_j^V ;
- $\{2^{-j} \psi_D(x 2^{-j} - n, y 2^{-j} - m), (n, m) \in \mathbb{Z}^2\}$ base orthonormale de W_j^D .

Ces trois types d'ondelettes ($\psi_H(x, y)$, $\psi_V(x, y)$ et $\psi_D(x, y)$) permettent ainsi d'extraire les détails d'une image à différentes échelles et orientations. Les transformées Fourier de $\varphi(t)$ et $(\psi_I(t))_{i \in \{H, V, D\}}$ sont respectivement de type passe-bas et passe-bande ce qui implique des propriétés de sélectivité

fréquentielles des ondelettes bidimensionnelles. Les ondelettes $\psi_H(x, y)$ vont mettre en évidence les détails de basses fréquences horizontales (suivant x) et hautes fréquences verticales (suivant y), ce qui correspond à des contours horizontaux. De même, $\psi_V(x, y)$ permet d'extraire les contours verticaux à savoir les hautes fréquences horizontales et les basses fréquences verticales. Enfin, $\psi_D(x, y)$ récupère les composantes résiduelles, qui sont de hautes fréquences, suivant les deux dimensions de l'espace et nommées détails diagonaux.

La projection d'une image $s(x, y)$ sur le sous-espace V_j^2 s'interprète comme une approximation de cette image à la résolution 2^{-j} et est donnée par la formule :

$$\text{proj}_{V_j^2}[s(x, y)] = \sum_{n=-\infty}^{+\infty} \sum_{m=-\infty}^{+\infty} a_j[n, m] 2^{-j} \varphi(x 2^{-j} - n) \varphi(y 2^{-j} - m) \quad (\text{C.33})$$

avec

$$a_j[n, m] = \langle \langle s(x, y), 2^{-j} \varphi(x 2^{-j} - n) \varphi(y 2^{-j} - m) \rangle \rangle \quad (\text{C.34})$$

où $\langle \langle \cdot, \cdot \rangle \rangle$ désigne le produit scalaire de $L^2(\mathbb{R})$.

Toutefois, comme en monodimensionnel C.1.6, il est possible de faire les approximations de manière récursive en exploitant les filtres de l'AMR :

$$a_j[n, m] = \sum_{k=-\infty}^{+\infty} \sum_{l=-\infty}^{+\infty} a_{j-1}[k, l] h_0^*[k - 2n] h_0^*[l - 2m] \quad (\text{C.35})$$

Le passage entre deux approximations successives se fait donc par filtrage passe-bas 2D de réponse impulsionnelle $(h_0^*[-k] h_0^*[-l])_{(k,l) \in \mathbb{Z}^2}$ suivie d'une décimation d'un facteur 2 suivant les lignes et les colonnes de l'image comme on peut le voir sur la figure C.7.

Enfin, de la même façon, on obtient les projections sur les sous-espaces $(W_j^i)_{i \in \{H, V, D\}}$:

$$\text{proj}_{V_j^i}[s(x, y)] = \sum_{n=-\infty}^{+\infty} \sum_{m=-\infty}^{+\infty} c_j^i[n, m] 2^{-j} \psi_i(x 2^{-j} - n, y 2^{-j} - m) \quad (\text{C.36})$$

avec

$$c_j^i[n, m] = \langle \langle s(x, y), 2^{-j} \psi_i^*(x 2^{-j} - n, y 2^{-j} - m) \rangle \rangle \quad (\text{C.37})$$

On déduit en le schéma de décomposition de la figure C.7.

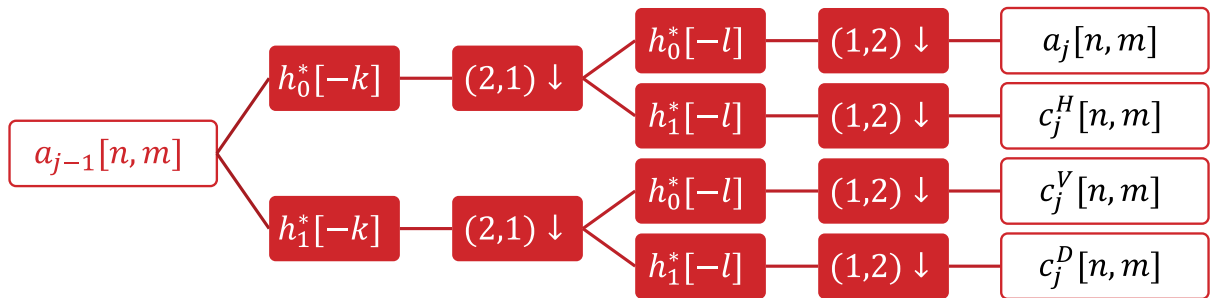


FIGURE C.7 – Structure d'un banc de filtres d'analyse en ondelettes séparables 2D.

La décomposition d'une image pour 3 niveaux de résolution ($j \in \{1, 2, 3\}$) se présente sous la forme du schéma de la figure C.8. Au premier niveau de résolution $j = 1$, l'image d'approximation est placée en haut à gauche, les contours horizontaux en haut à droite, les verticaux en bas à gauche et les diagonaux en bas à droite. Puis, la décomposition au niveau $j = 2$ se déduit de l'image d'approximation du niveau $j - 1$ qui est alors remplacée par la nouvelle image d'approximation, les contours horizontaux, les verticaux et les diagonaux. Cette opération peut être répétée autant de fois que nécessaire (dans la limite de la taille de l'image qui est divisée par 4 à chaque niveau de résolution). Une décomposition d'une image en ondelettes de Haar est illustrée sur la figure C.9.

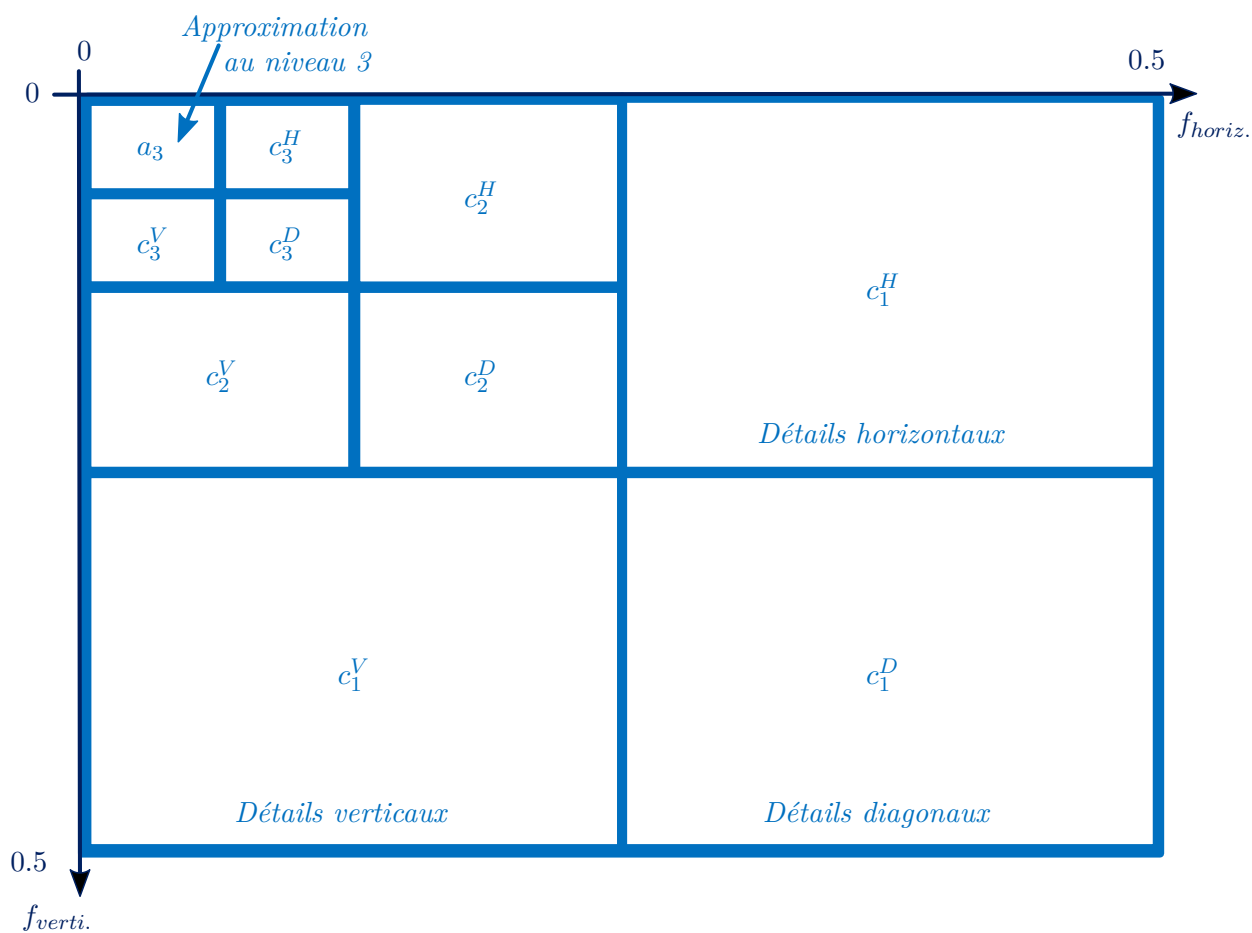
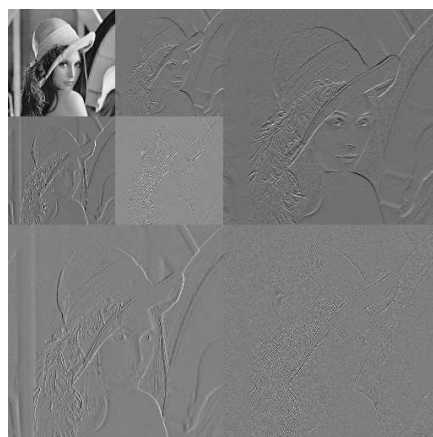


FIGURE C.8 – Représentation d’une AMR en ondelettes séparables sur 3 niveaux de résolutions.



(a) Image originale



(b) Décomposition de Haar

FIGURE C.9 – AMR en ondelettes séparables sur 3 niveaux de résolutions d’une image.

La reconstruction d'une image à partir de sa transformée en ondelettes se fait en cascasant de façon séparable des bancs de filtres de synthèse monodimensionnels. En image, on peut également généraliser aux ondelettes M -bandes et aux paquets d'ondelettes. On obtiendrait alors des représentations telles que celles exposées sur la figure C.10 où l'on peut apercevoir une décomposition en ondelettes 3-bandes à gauche et une décomposition en paquets d'ondelettes à droite.

c_{200}	c_{210}	c_{220}	$c_{1,1,0}$	$c_{1,2,0}$
c_{201}	c_{211}	c_{221}		
c_{202}	c_{212}	c_{222}		
$c_{1,0,1}$			$c_{1,1,1}$	$c_{1,2,1}$
$c_{1,0,2}$			$c_{1,1,2}$	$c_{1,2,2}$

(a) Ondelettes de Haar

$c_{2,0,0}$	$c_{2,1,0}$	$c_{2,2,0}$	$c_{2,3,0}$
$c_{1,0,0}$		$c_{1,1,0}$	
$c_{2,0,1}$	$c_{2,1,1}$	$c_{2,2,1}$	$c_{2,3,1}$
$c_{2,0,2}$	$c_{2,1,2}$	$c_{2,2,2}$	$c_{2,3,2}$
$c_{1,0,1}$		$c_{1,1,1}$	
$c_{2,0,3}$	$c_{2,1,3}$	$c_{2,2,3}$	$c_{2,3,3}$

(b) Ondelettes du chapeau mexicain

FIGURE C.10 – Représentation d'une AMR en ondelettes 3-bandes et en paquet d'ondelettes pour 2 niveaux de résolutions.

C.2.2 Analyse en ondelettes non séparables

Contrairement à la section précédente, cette section considère des bancs de filtres construits directement en deux dimensions [18], ce qui implique que les ondelettes sont dans $L^2(\mathbb{C})$. On parle alors d'ondelettes non séparables. A partir d'une ondelette $\psi(\vec{x})$, la famille d'ondelette est définie, pour tout vecteur \vec{x} de \mathbb{R}^2 (ou \mathbb{C}), par dilatation/contraction, rotation et translation :

$$\psi_{(a, \vec{b}, \theta)}(\vec{x}) = \frac{1}{a} \psi \left(R^{-\theta} \left(\frac{\vec{x} - \vec{b}}{a} \right) \right) \quad (\text{C.38})$$

avec \vec{b} le vecteur translation réel (ou complexe), a une échelle positive et $R^{-\theta}$ la rotation d'angle θ définie par la matrice $R^{-\theta} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix}$.

Si l'ondelette mère ψ est parfaitement isotrope, c'est-à-dire que, quel que soit la direction θ , on a $\psi(R^{-\theta}\vec{x}) = \psi(\vec{x})$ pour tout \vec{x} . Le paramètre θ n'a alors aucun sens et l'ondelette est dite omnidirectionnelle. Dans le cas contraire, on parlera d'ondelettes directionnelles car elles ont la propriété d'« observer » les caractéristiques du signal dans une direction donnée tout en étant « aveugle » dans les autres directions. L'ondelette directionnelle la plus courante est celle de Morlet (voir équation (C.8)). Par exemple sur la figure C.11, le caractère directionnel de l'ondelette de Morlet lui permet de distinguer les contours verticaux, horizontaux et diagonaux. Par contre, l'ondelette du chapeau mexicain, qui est omnidirectionnelle, affiche l'ensemble des contours sur sa transformée pour chacune des échelles.

La transformée en ondelette directionnelle d'une image $s(\vec{x}) \in L^2(\mathbb{R}^2)$ est définie par :

$$W_f(a, \vec{x}, \theta) = \int \int_{\mathbb{R}^2} s(\vec{x}) \psi_{(a, \vec{x}, \theta)}^*(\vec{x}) d\vec{x} \quad (\text{C.39})$$

Il est également possible de reconstruire l'image à partir de la transformée en ondelettes directionnelles.

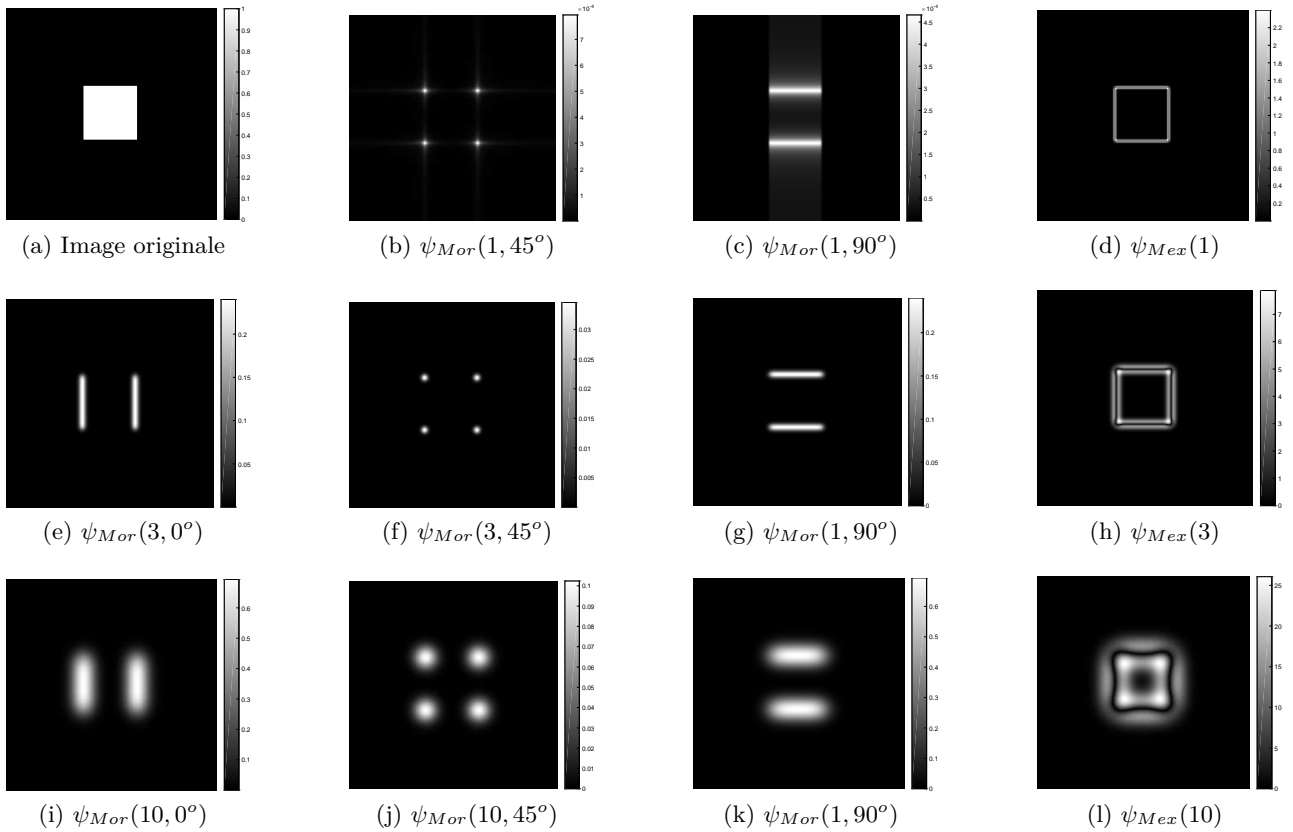


FIGURE C.11 – Transformée en ondelettes 2D directionnelles (Morlet ψ_{Mor}) et omnidirectionnelles (chapeau mexicain ψ_{Mex}) d'un carré sous plusieurs échelles $\{1, 3, 10\}$ et orientations $\{0^\circ, 45^\circ, 90^\circ\}$.

Les ondelettes directionnelles sont localisées en fréquence. La figure C.12 illustre un exemple de balayage fréquentiel par des ondelettes filles qui évolue en fonction de l'angle θ et du facteur d'échelle a . Si le carré gris central représente le support de la transformée de Fourier de l'ondelette mère, les deux autres carrés représentent les supports des ondelettes filles et illustrent l'effet de a et θ dans le plan fréquentiel. Le centre de la fonction subit une rotation d'angle θ et une homothétie de rapport $1/a$, et « l'aire » du support diminue lorsque a augmente.

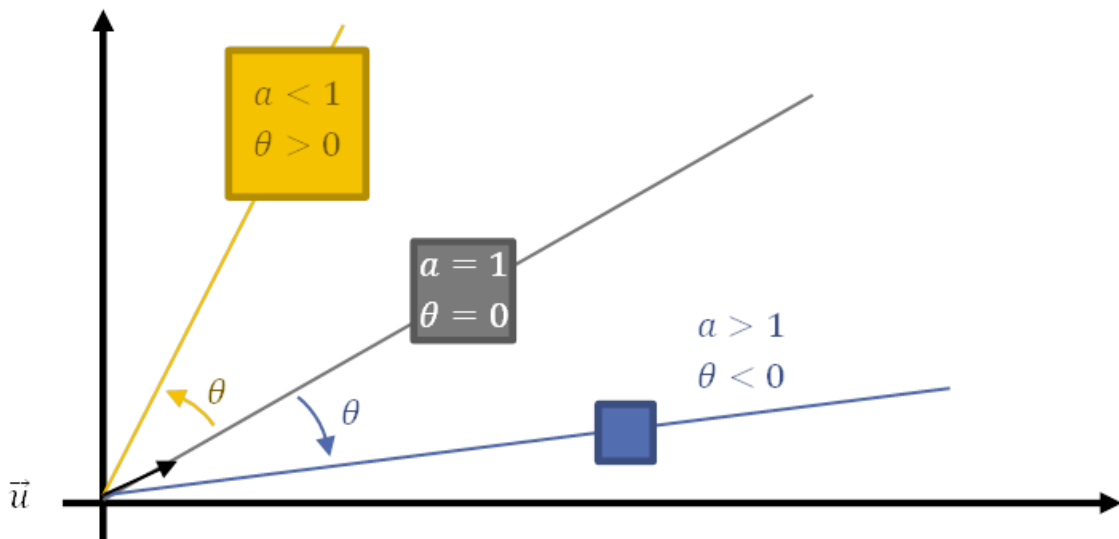


FIGURE C.12 – Balayage fréquentiel par des ondelettes 2D en fonction de l'angle θ et du facteur d'échelle a (inspiré de [18]).

Annexe D

Complément sur les statistiques d'ordre supérieur

Introduction

Dans le but de caractériser la structure d'un thrombus veineux, nous proposons une approche d'extraction d'attributs de nos images de thrombi basée sur les statistiques d'ordre supérieur. Avant d'envisager leur application à des images, cet annexe propose, en section D.1, une introduction aux statistiques d'ordre supérieur, et plus particulièrement aux concepts de moments et de cumulants. La section D.2 est consacrée aux propriétés d'ordre supérieur des signaux aléatoires dans le domaine temporel : la multicorrélation. Ensuite, l'extension des multicorrélations des signaux temporels aux images et la construction des descripteurs statistiques de nos images de thrombi font l'objet de la section 4.2 du chapitre 4.

D.1 Introduction aux statistiques d'ordre supérieur

Dans cette section, nous allons introduire un certain nombre de concepts de la théorie des probabilités et plus particulièrement les moments et les cumulants d'ordre supérieur. Dans cette section sur les statistiques d'ordre supérieur, **le nombre complexe $i = \sqrt{-1}$ sera exceptionnellement noté j** pour qu'il ne soit pas confondu avec l'indice i utilisé tout au long de cette section.

D.1.1 Définitions préliminaires

Cette sous-section définit certaines notions de base en probabilité et statistique nécessaires pour pouvoir aborder ensuite les statistiques d'ordre supérieur.

D.1.1.1 Évènement et expérience aléatoire

Les phénomènes physiques et les instruments d'observation comme les échographes sont souvent les sources de résultats imprévisibles sous certaines limites [107]. Une *expérience aléatoire* est une expérience qui, répétée dans les mêmes conditions, ne donnera pas forcément le même résultat. L'ensemble des résultats que peut fournir cette expérience constitue l'*espace des réalisations* Ω .

Par exemple, lorsque nous lançons un dé, on réalise une expérience aléatoire dont l'espace des réalisations $\Omega = \{1, 2, 3, 4, 5, 6\}$ est de dimension finie [107]. Chaque élément ω d'un espace des réalisations Ω dénombrable est appelé *événement élémentaire*. Un sous-ensemble de Ω peut être un événement composé (plusieurs événements) ou simple (un seul). On dit qu'un événement $A \in \Omega$ est *réalisé* si le résultat ω de l'expérience aléatoire appartient à A .

D.1.1.2 Probabilité d'un évènement

Une probabilité est une fonction qui associe à un évènement un nombre compris entre 0 et 1 [107]. Plus ce nombre est élevé, plus il est vraisemblable que l'évènement se réalise. Dans la littérature, il existe trois manières de définir la probabilité d'un évènement : classique, fréquentiste et axiomatique [120].

La définition classique (ou de Laplace) est donnée par [107] :

$$P(A) = \frac{N_A}{N} \quad (\text{D.1})$$

où N_A est le nombre d'évènements qui constituent A et N le cardinal de Ω (*i.e.* le nombre de cas total). Toutefois, cette définition est limitée puisqu'elle implique que Ω soit dénombrable et que les évènements élémentaires soient équiprobables (un dé pipé ne pourrait donc pas être modélisé par exemple).

La définition fréquentiste s'appuie sur la formulation de la loi des grands nombres pour définir une probabilité [120] :

$$P(A) = \lim_{N \rightarrow \infty} \frac{N_A}{N} \quad (\text{D.2})$$

où N est le nombre de fois qu'a été réalisée l'expérience aléatoire, et N_A est le nombre de fois que l'évènement A s'est produit parmi celles-ci. Cette définition a l'avantage d'être robuste face à un espace des réalisations infini et face à des évènements aléatoires non équiprobables, mais le passage à la limite conduit à des calculs lourds dans les cas simples où la définition classique est valide.

La troisième définition repose sur trois axiomes [120]. Une fonction P d'un ensemble d'évènements d'un espace de réalisations Ω est une probabilité si :

1. Pour tout évènement A , $P(A) \geq 0$;
2. $P(\Omega) = 1$;
3. Si $A \cap B = \emptyset$ alors $P(A \cup B) = P(A) + P(B)$ c'est-à-dire, que si A et B sont disjoints, alors la probabilité que A ou B arrive est égale à la somme entre la probabilité de A et celle de B .

Cette fonction P est appelée *loi* de probabilité. Pour une expérience aléatoire X , la probabilité de l'évènement ω est souvent notée $P(X = \omega)$ ou $P_X(\omega)$.

D.1.1.3 Variable aléatoire et fonction de répartition

Une expérience aléatoire produisant un nombre est appelée une *variable aléatoire* [107]. Son espace des réalisations est donc un sous-ensemble de \mathbb{R} ou de \mathbb{C} . Dans un espace probabiliste (Ω, P) , une variable aléatoire est l'image $X(\omega)$ d'une application X de l'espace des épreuves Ω dans \mathbb{R} (ou \mathbb{C}) :

$$\begin{aligned} X : \Omega &\rightarrow \mathbb{R} \\ \omega &\mapsto X(\omega) \end{aligned} \quad (\text{D.3})$$

Une variable aléatoire $X(\omega)$, notée X en pratique, est dite *discrète* lorsque son espace des réalisations $\Omega = \{\omega_1, \omega_2, \dots\}$ est fini ou dénombrable¹. Sa loi est définie par les probabilités des évènements élémentaires ω_i :

$$P(X = \omega_i) = P_X(\omega_i) = P_i \quad (\text{D.4})$$

Dans le cas contraire, si l'espace des réalisations est indénombrable, la variable est dite *continue* et sa mesure de probabilité est une mesure de densité, appelée *densité de probabilité* ou *fonction de*

1. Un ensemble E est dit dénombrable s'il existe une bijection entre l'ensemble N des entiers naturels et E [107].

densité. La densité de probabilité $f_X(x)$ d'une variable aléatoire X est une application de $\mathbb{R} \rightarrow \mathbb{R}^+$ (où \mathbb{R}^+ désigne l'ensemble des nombres réels positifs) vérifiant les propriétés suivantes :

$$f_X(x) \geq 0, \forall x \in \mathbb{R} \quad (\text{D.5})$$

$$\int_{-\infty}^{\infty} f_X(x) dx = 1 \quad (\text{D.6})$$

Dans le cas général, la probabilité d'un évènement A est définie par :

$$P_X(A) = P(X \in A) = \int_A f_X(x) dx \quad (\text{D.7})$$

Si X est une variable aléatoire réelle, sa *fonction de répartition* F_X est définie sur \mathbb{R} par [107] :

$$F_X(x) = P(X \leq x) = P(X \in]-\infty, x]) = \int_{-\infty}^x f_X(t) dt \quad (\text{D.8})$$

La fonction de répartition ou la densité de probabilité d'une variable aléatoire permet de définir la variable aléatoire.

D.1.1.4 Espérance mathématique

L'*espérance* est la valeur que l'on s'attend à trouver, en moyenne, si l'on répète un grand nombre de fois la même expérience aléatoire. Pour une grandeur physique comme l'intensité $I(t)$, l'espérance représente son énergie continue [107]. D'un point de vue mathématique, elle est définie par :

$$E[X] = \int_{\Omega} X(\omega) dP_X(\omega) = \int_{\mathbb{R}} x dP_X(x) \quad (\text{D.9})$$

L'espérance d'une variable continue X est :

$$E[X] = \int_{-\infty}^{+\infty} x f_X(x) dx \quad (\text{D.10})$$

Pour une variable discrète X , l'espérance est :

$$E[X] = \sum_i \omega_i P_X(\omega_i) = \sum_i \omega_i P_i \quad (\text{D.11})$$

L'espérance d'une variable aléatoire est également son moment d'ordre un, qui sera noté $m_1[X]$. Une variable aléatoire d'espérance nulle est dite *centrée*.

Si l'on considère X et Y deux variables aléatoires et a et b deux variables certaines (nombres), alors l'espérance vérifie les propriétés suivantes [107] :

$$E[a] = a \quad (\text{D.12})$$

$$E[aX + bY] = aE[X] + bE[Y] \quad (\text{D.13})$$

$$|E[X]| \leq E[|X|] \quad (\text{D.14})$$

Les équations (D.13) et (D.14) indiquent, respectivement, que l'espérance est une application linéaire.

D.1.1.5 Variance et moment d'ordre deux

L'espérance mathématique d'une variable aléatoire peut être interprétée comme le centre de gravité de son espace des réalisations. Toutefois, elle ne présage rien sur la forme de la densité de probabilité. Pour caractériser une variable aléatoire, il faut souvent utiliser d'autres quantités statistiques. Parmi celles-ci, le *moment d'ordre deux* qui permet de mesurer la dispersion de la dite-variable [107].

Le moment d'ordre deux d'une variable aléatoire continue X est donné par :

$$m_2[X] = E[X^2] = \int_{-\infty}^{+\infty} x^2 f_X(x) dx \quad (D.15)$$

où $f_X(x)$ est la densité de probabilité de X .

Dans le cas d'une variable X discrète avec une loi $P(X = \omega_i) = P_i$, le moment d'ordre deux est :

$$m_2[X] = E[X^2] = \sum_i \omega_i^2 P_X(\omega_i) = \sum_i \omega_i^2 P_i \quad (D.16)$$

En pratique, il est d'usage d'utiliser les moments centrés plutôt que les moments :

$$\mu_2 = \sigma_X^2 = E[(X - E[X])^2] = m_2[X] - (m_1[X])^2 \geq 0 \quad (D.17)$$

Le moment centré d'ordre deux, μ_2 , représente la *variance* de la variable aléatoire X et peut être interprétée comme l'énergie dynamique de la variable X : elle caractérise la dispersion des réalisations de la variable X autour de sa valeur moyenne [107]. Une variable aléatoire de variance unitaire ($\mu_2 = 1$) est dite *réduite*.

Si X est une variable aléatoire réelle, la racine de sa variance σ_X^2 correspond à l'*écart-type* σ_X de la variable X . Si X représente une grandeur physique, l'écart-type σ_X possèdera la même unité. L'Analyse en Composantes Principales (ACP), que nous avons abordée en section 2.3.2, utilise notamment la variance pour déterminer un nouvel espace de dimension réduite qui maximise la dispersion des variables.

D.1.1.6 Loi normale ou gaussienne

A présent, nous allons voir un exemple de la loi usuelle de probabilité, la loi *gaussienne* [107]. On parle également de loi *normale* car elle est souvent utilisée en première approximation pour modéliser le bruit additif (bruit thermique, bruit de rayonnement, *etc.*) dans de nombreuses applications.

Une variable aléatoire gaussienne de moyenne m et d'écart type σ possède une densité de probabilité définie par :

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right) \quad (D.18)$$

avec $E[X] = m$ et $\mu_2 = \sigma^2$. En pratique, elle est notée $\mathcal{N}(m, \sigma^2)$. On dit que la variable aléatoire X suit une loi gaussienne centrée réduite si $m = 0$ et $\sigma = 1$. Elle est alors notée $\mathcal{N}(0, 1)$ et visible en figure D.1.

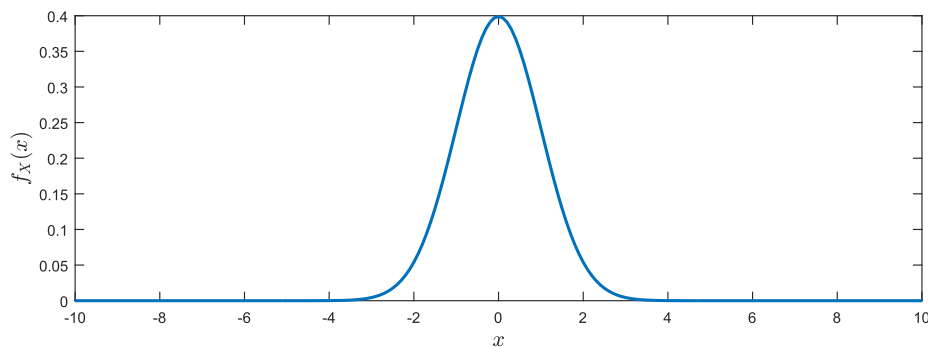


FIGURE D.1 – Densité de probabilité d'une variable gaussienne centrée et réduite

La fonction de répartition d'une variable gaussienne est :

$$F_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x \exp\left(-\frac{(u-m)^2}{2\sigma^2}\right) du \quad (\text{D.19})$$

Dans ce mémoire, cette loi est notamment évoquée aux sections 2.4.2.3 et 4.1.2 pour définir les notions de similarité gaussienne (pour la classification spectrale) et de filtre passe-bas gaussien (pour le *scattering operator*).

D.1.2 Statistiques d'ordre supérieur

En pratique, nous nous retrouvons souvent à traiter des variables aléatoires dont la loi de probabilité est inconnue [132]. C'est notamment le cas si nous considérons nos images comme des variables aléatoires bidimensionnelles. Pour estimer les densités de probabilité, la littérature propose divers estimateurs mais ces derniers sont souvent difficiles à mettre en place et lourds en temps de calcul. Toutefois, il existe des algorithmes qui n'utilisent pas directement les densités de probabilité des variables aléatoires pour étudier un phénomène aléatoire. Ces algorithmes s'intéressent à d'autres quantités statistiques qui sont définies à partir des densités de probabilité et qui sont plus facilement estimables. Il y a notamment les moments ou les cumulants d'ordre supérieur des variables aléatoires [132]. Certains estimateurs donnent des valeurs approchées de ces moments à partir d'un nombre limité de tirages aléatoires d'une variable aléatoire inconnue et, fournissent ainsi des outils d'extrapolation de son comportement.

D.1.2.1 Moments d'ordre supérieur

Soit X une variable aléatoire continue de densité de probabilité $f_X(x)$, son *moment d'ordre q* est défini par :

$$m_q[X] = E[X^q] = \int_{-\infty}^{+\infty} x^q f_X(x) dx \quad (\text{D.20})$$

Le moment d'ordre q d'une variable discrète X qui suit une loi de probabilité $P(X = x_i) = P_i$ est donné par :

$$m_q[X] = E[X^q] = \sum_i x_i^q P_X(x_i) = \sum_i x_i^q P_i \quad (\text{D.21})$$

Comme évoqué précédemment, il est plus fréquent de considérer les *moments centrés d'ordre q* :

$$\mu_q[X] = E[(X - E[X])^q] \quad (\text{D.22})$$

Si la variable aléatoire X a une distribution de probabilité paire, et que l'intégrale $\int_{-\infty}^{+\infty} x^{2q+1} f_X(x) dx$ (ou la somme $\sum_i x_i^q P_i$) converge, alors les moments d'ordre impair $m_{2q+1}[X]$ sont nuls. De plus, comme l'espérance $E[X]$ est nulle, on a aussi $m_q[X] = \mu_q[X]$. Pour une loi normale, les moments centrés d'ordre supérieur à deux sont nuls.

D.1.2.2 Fonctions caractéristiques

Les fonctions caractéristiques permettent, comme la densité de probabilité, de décrire complètement une variable aléatoire [50]. La *première fonction caractéristique* d'une variable X est :

$$\Phi_X(t) = E[e^{jXt}] \quad (\text{D.23})$$

où $j = \sqrt{-1}$.

La première fonction caractéristique d'une variable continue est donc :

$$\Phi_X(t) = \int_{-\infty}^{+\infty} e^{jxt} f_X(x) dx \quad (\text{D.24})$$

alors que celle d'une variable discrète est :

$$\Phi_X(t) = \sum_i e^{jx_i t} P_X(x_i) \quad (\text{D.25})$$

La seconde fonction caractéristique est donnée par :

$$\Psi_X(t) = \ln(\Phi_X(t)) \quad (\text{D.26})$$

Nous pouvons constater que la première fonction caractéristique n'est autre qu'une transformée de Fourier de sa mesure de probabilité, et il est donc possible de retrouver la densité de probabilité en calculant la transformée de Fourier inverse :

$$f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-jxt} \Phi_X(t) dt \quad (\text{D.27})$$

En outre, si la première fonction caractéristique est infiniment dérivable en zéro, alors les moments d'ordre q peuvent se déduire de son développement en série entière :

$$m_q[X] = E[X^q] = (-1)^q \left. \frac{\partial^q \Phi_X(t)}{\partial t^q} \right|_{t=0} \quad (\text{D.28})$$

Par conséquent, la connaissance de tous les moments d'une variable entraîne la connaissance de sa première fonction caractéristique par son développement en série entière, et donc, la connaissance de sa loi d'évolution [50].

D.1.2.3 Cumulants d'ordre supérieur

Dans la pratique, les statistiques d'ordre supérieur sous la forme de cumulants sont souvent privilégiées par rapport à celles sous forme de moments [107, 115]. Les propriétés algébriques des cumulants rendent leur utilisation particulièrement intéressante car elles permettent de simplifier les calculs numériques et d'améliorer les performances de certains algorithmes de traitement du signal.

Les *cumulants d'ordre q* proviennent du développement en série entière de la seconde fonction caractéristique $\Psi_X(t)$:

$$c_q = c[X^q] = (-1)^q \left. \frac{\partial^q \Psi_X(t)}{\partial t^q} \right|_{t=0} \quad (\text{D.29})$$

D.1.2.4 Propriétés entre moments et cumulants

Pour une variable aléatoire X , le cumulant d'ordre q de X est lié aux moments d'ordres inférieurs ou égaux à q par la formule dite de Leonov et Shiryayev [94] :

$$c_q[X] = \sum_{p=1}^q (-1)^{p-1} (p-1)! m_{\nu_1} m_{\nu_2} \dots m_{\nu_p} \quad (\text{D.30})$$

où la somme porte sur toutes les permutations $\{\nu_1, \nu_2, \dots, \nu_p : 1 \leq p \leq q\}$ telles que les nombres ν_i appartiennent à l'ensemble $\{1, \dots, q-p+1\}$ et telles que la somme de ces nombres $\sum_{i=1}^p \nu_i$ soit égale à q .

Pour l'ordre quatre, les combinaisons possibles des indices i, j, k et l en p ensemble(s) sont :

$$\mathbf{p} = \mathbf{1} : i|jkl$$

$$\mathbf{p} = \mathbf{2} : i|jkl ; j|ikl ; k|ijl ; l|ijk \\ ij|kl ; ik|jl ; il|jk$$

$$\mathbf{p} = \mathbf{3} : i|j|kl ; i|k|jl ; i|l|jk ; j|k|il ; j|l|ik ; k|l|ij$$

$$\mathbf{p} = \mathbf{4} : i|j|k|l$$

Pour $p = 2$, il y a donc sept manières de créer deux ensembles à partir des quatre indices : par exemple, $i|jkl$ indique qu'il faut calculer le produit $m_1[X_i]m_3[X_j, X_k, X_l]$ pour obtenir le cumulants croisé d'ordre quatre, $c_4[X_i, X_j, X_k, X_l]$, des variables X_i, X_j, X_k et X_l .

A partir de l'équation (D.30) et des différentes combinaisons énumérées ci-dessus, le cumulants d'ordre quatre est donné par :

$$\begin{aligned}
 c_4 &= (-1)^0 (0)! m_4 \\
 &+ (-1)^1 (1)! (4m_1m_3 + 3m_2^2) \\
 &+ (-1)^2 (2)! 6m_1^2m_2 \\
 &+ (-1)^3 (3)! m_1^4 \\
 &= m_4 - 4m_1m_3 - 3m_2^2 + 12m_1^2m_2 - 6m_1^4
 \end{aligned} \tag{D.31}$$

Cette équation peut être simplifiée si la variable aléatoire est centrée :

$$c_4 = \mu_4 - 3\mu_2^2 \tag{D.32}$$

Les cumulants d'ordre q peuvent également s'obtenir à partir des moments d'ordre inférieur ou égal à q par la relation de récurrence [107] :

$$c_q = m_q - \sum_{k=1}^{q-1} \binom{q-1}{k-1} c_k m_{q-k} \tag{D.33}$$

où $\binom{q}{k} = \frac{q!}{k!(q-k)!}$ sont les coefficients binomiaux.

A partir de cette relation, on peut en déduire les cumulants d'ordres trois, cinq et six :

$$c_3 = m_3 - 3m_2m_1 + 2m_1^3 = \mu_3 \tag{D.34}$$

$$\begin{aligned}
 c_5 &= m_5 - 5m_4m_1 - 10m_3m_2 + 20m_3m_1^2 + 30m_2^2m_1 - 60m_2m_1^3 + 24m_1^5 \\
 &= \mu_5 - 10\mu_3\mu_2
 \end{aligned} \tag{D.35}$$

$$\begin{aligned}
 c_6 &= m_6 - 6m_5m_1 - 15m_4m_2 - 10m_3^2 + 30m_4m_1^2 + 120m_3m_2m_1 + 30m_2^3 \\
 &- 120m_3m_1^3 - 270m_2^2m_1^2 + 120m_2m_1^4 - 120m_1^6 \\
 &= \mu_6 - 15\mu_4\mu_2 - 10\mu_3^2 + 30\mu_2^3
 \end{aligned} \tag{D.36}$$

D.1.2.5 Interprétation des cumulants

Pour toute variable aléatoire X , le cumulants d'ordre un, $c_1[X]$, est l'espérance de X . Le cumulants $c_2[X]$ correspond à la variance et s'identifie à l'énergie dynamique de X . Le cumulants $c_3[X]$ désigne la dissymétrie de la densité de probabilité de X . Le coefficient adimensionnel ξ_X d'une variable centrée est appelé *coefficient d'asymétrie* (*skewness* en anglais) et est donné par :

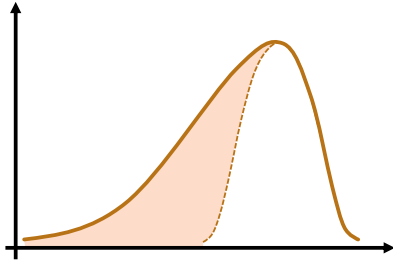
$$\xi_X = \frac{\mu_3[X]}{\sigma_X^3} \tag{D.37}$$

Un coefficient positif de l'asymétrie indique un étalement vers les valeurs positives de X comme nous pouvons le voir sur la figure D.2.b.

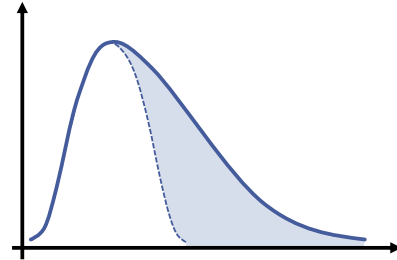
Ensuite, le cumulants d'ordre quatre, $c_4[X]$, caractérise l'aplatissement de la densité de probabilité de X par rapport à la loi normale. Le *coefficient d'aplatissement* (ou *kurtosis*) [109, 110] d'une variable est définie par :

$$\kappa_X = \frac{c_4[X]}{\sigma_X^4} \tag{D.38}$$

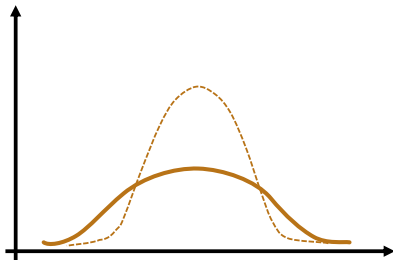
Un coefficient d'aplatissement positif indique que la densité de probabilité est plus tassée, et que ses branches infinies sont plus épaisses, que la loi normale (voir figure D.2.d). Il est à noter que cette interprétation du *kurtosis* est uniquement valable dans le cas particulier où la densité de probabilité de la variable aléatoire X suit une loi unimodale² qui ne possède que deux intersections (dans \mathbb{R}) avec la loi normale.



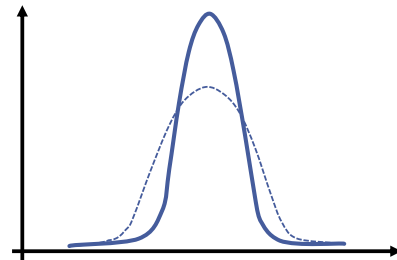
(a) Un coefficient d'asymétrie négatif indique une distribution décalée à gauche de la médiane



(b) Un coefficient d'asymétrie positif suggère une distribution décalée à droite de la médiane



(c) Un coefficient d'aplatissement négatif signifie que la distribution est platykurtique par rapport à la distribution normale



(d) Un coefficient d'aplatissement positif correspond à une distribution leptokurtique par rapport à la distribution normale

FIGURE D.2 – Interprétation des cumulants d'ordres trois et quatre normalisés appelés respectivement coefficients d'asymétrie et d'aplatissement : cette interprétation est uniquement valable dans le cas particulier d'une densité de probabilité unimodale avec seulement deux intersections avec la loi normale [109, 110].

En pratique, les statistiques d'ordre supérieur à quatre sont peu utilisées. Contrairement aux ordres précédents, les ordres supérieurs à quatre ne sont pas interprétables physiquement et sont d'énergie plus faibles.

D.1.3 Couple de variables aléatoires

Dans cette sous-section, nous allons considérer un couple de variables aléatoires réelles et définir les notions de la section précédente pour ce couple. Dorénavant, nous supposons que les variables aléatoires sont continues bien qu'il soit possible de généraliser aux variables discrètes à partir des définitions des sous-sections D.1.1 et D.1.2.

D.1.3.1 Lois de probabilité conjointes et marginales

La fonction de répartition, définie à la sous-section D.1.1.3, peut se généraliser à un couple (X, Y) de variables aléatoires réelles :

$$F_{X,Y}(x, y) = P(X \leq x \text{ et } Y \leq y) \quad (\text{D.39})$$

2. La densité de probabilité $f_X(x)$ d'une variable aléatoire X est dite *unimodale* s'il existe un intervalle $[a, b]$, éventuellement réduit à un point, tel que $f_X(x)$ soit croissante sur $] -\infty, a[$, constante sur $[a, b]$ et décroissante sur $]b, +\infty[$ [6].

On parle généralement de *fonction de répartition conjointe* et, parmi ses propriétés, on a X et Y indépendantes si et seulement si :

$$F_{X,Y}(x, y) = F_X(x)F_Y(y) \quad (\text{D.40})$$

On dit que le couple (X, Y) est continu à densité si $F_{X,Y}(x, y)$ est dérivable par rapport à chacune de ses variables, c'est-à-dire que les valeurs du couple (X, Y) sont dans une région au voisinage de (x, y) [107]. La *densité conjointe* est alors :

$$f_{X,Y}(x, y) = \frac{\partial^2 F_{X,Y}(x, y)}{\partial x \partial y} \quad (\text{D.41})$$

À partir de la loi conjointe du couple (X, Y) , il est possible de retrouver la loi de chacune des variables (*i.e.* les lois marginales). En revanche, l'inverse n'est pas vrai : connaître la loi de X et celle de Y ne permet pas de construire leur loi conjointe (sauf si X et Y sont indépendantes).

Les lois de probabilité séparées, dites *marginales* de X et de Y sont définies par :

$$\mathbf{X} : \quad f_X(x) = \int_{-\infty}^{+\infty} f_{X,Y}(x, y) dy \quad \text{et} \quad F_X(x) = F_{X,Y}(x, +\infty) \quad (\text{D.42})$$

$$\mathbf{Y} : \quad f_Y(y) = \int_{-\infty}^{+\infty} f_{X,Y}(x, y) dx \quad \text{et} \quad F_Y(y) = F_{X,Y}(+\infty, y) \quad (\text{D.43})$$

D.1.3.2 Espérance et moments d'ordre deux

L'espérance du couple de variables aléatoires (X, Y) est constituée de ses moments d'ordre un :

$$\mathbf{E}[(X, Y)] = (\mathbf{E}[X], \mathbf{E}[Y]) \quad (\text{D.44})$$

De manière plus générale, le moment d'ordre p et q du couple (X, Y) est $\mathbf{E}[X^p, Y^q]$ et le moment centré est $\mathbf{E}[X_c^p, Y_c^q]$ où $X_c = X - \mathbf{E}[X]$ et $Y_c = Y - \mathbf{E}[Y]$ désignent les variables aléatoires centrées obtenues par translation des variables X et Y .

À l'ordre deux, nous obtenons donc trois moments : $\mathbf{E}[X^2]$, $\mathbf{E}[Y^2]$ (les moments d'ordre deux de X et Y) et $\mathbf{E}[XY]$ qui est le moment croisé d'ordre deux aussi appelé *corrélacion* entre X et Y . Si la corrélation est nulle, cela indique que les variables X et Y sont orthogonales.

De la même façon, il y a trois moments centrés d'ordre deux : les variances $\mathbf{E}[X_c^2]$, $\mathbf{E}[Y_c^2]$ et $\mathbf{E}[X_c Y_c]$ qui correspond à la *covariance* de X et Y , notée $\text{Cov}(X, Y)$. La covariance peut, comme pour la variance, être obtenue à partir des moments non centré :

$$\text{Cov}(X, Y) = \mathbf{E}[XY] - \mathbf{E}[X]\mathbf{E}[Y] \quad (\text{D.45})$$

La covariance de X et Y peut s'interpréter comme l'énergie dynamique croisée entre X et Y . La covariance normalisée par les écart-types de X et Y définit le *coefficient de corrélation* :

$$\rho = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y} \quad (\text{D.46})$$

Le coefficient de corrélation ρ est compris entre -1 et 1 : plus sa valeur absolue $|\rho|$ est importante, plus la relation de dépendance entre X et Y est forte. Lorsque ce coefficient est nul, on dit que les variables X et Y sont *décorrélées*. Deux variables indépendantes sont décorrélées mais l'inverse n'est pas vrai.

D.1.3.3 Fonctions caractéristiques et cumulants

Les fonctions caractéristiques d'une variable aléatoire s'étendent à un couple de variables aléatoires (X, Y) à l'instar de la transformée de Fourier qui s'étend aux fonctions de \mathbb{R}^2 :

$$\Phi_{X,Y}(u, v) = \mathbb{E} \left[e^{j(uX+vY)} \right] \quad (\text{D.47})$$

$$\Psi_{X,Y}(u, v) = \ln(\Phi_{X,Y}(u, v)) \quad (\text{D.48})$$

Les équations (D.28) et (D.29) s'étendent également en :

$$m_{p,q}[X, Y] = E[X^p Y^q] = (-1)^{p+q} \frac{\partial^{p+q} \Phi_{X,Y}(u, v)}{\partial u^p \partial v^q} \Big|_{(u,v)=(0,0)} \quad (\text{D.49})$$

$$c_{p,q}[X, Y] = c[\underbrace{X, \dots, X}_p \text{ fois}, \underbrace{Y, \dots, Y}_q \text{ fois}] = (-1)^{p+q} \frac{\partial^{p+q} \Psi_{X,Y}(u, v)}{\partial u^p \partial v^q} \Big|_{(u,v)=(0,0)} \quad (\text{D.50})$$

où $m_{p,q}[X, Y]$ et $c_{p,q}[X, Y]$ sont respectivement les moments et les cumulants croisés d'ordres (p, q) de (X, Y) .

Si X et Y sont deux variables indépendantes, alors on a les propriétés suivantes :

$$\Phi_{X,Y}(u, v) = \Phi_X(u) \Phi_Y(v) \quad (\text{D.51})$$

$$\Psi_{X,Y}(u, v) = \Psi_X(u) + \Psi_Y(v) \quad (\text{D.52})$$

D.1.4 Vecteurs aléatoires

Avant de s'attaquer aux statistiques d'ordre supérieur sur des images, nous allons d'abord les définir dans le cas de vecteurs aléatoires réels qui constitue une généralisation du cas d'un couple de variables aléatoires.

D.1.4.1 Lois de probabilité

Un vecteur aléatoire réel $X = (X_1, X_2, \dots, X_n)^T$ est un vecteur colonne de \mathbb{R}^n dont les composantes X_i sont des variables aléatoires. La densité de probabilité conjointe de ses composantes X_i est [107] :

$$f_X(x) = f_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n) \quad (\text{D.53})$$

Les composantes X_i d'un vecteur aléatoire sont indépendantes les unes des autres si et seulement si l'égalité suivante est vérifiée :

$$f_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n) = f_{X_1}(x_1) f_{X_2}(x_2) \dots f_{X_n}(x_n) \quad (\text{D.54})$$

La fonction de répartition du vecteur X est donnée par :

$$\begin{aligned} F_X(x) &= F_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n) \\ &= \mathbb{P}(X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n) \\ &= \int_{-\infty}^{x_n} \dots \int_{-\infty}^{x_2} \int_{-\infty}^{x_1} f_{X_1, X_2, \dots, X_n}(t_1, t_2, \dots, t_n) dt_1, dt_2, \dots, dt_n \end{aligned} \quad (\text{D.55})$$

D.1.4.2 Fonctions caractéristiques, moments et cumulants

De la même manière que la transformée de Fourier, les fonctions caractéristiques deviennent des fonctions de \mathbb{R}^n dans \mathbb{C} :

$$\Phi_X(u) = \mathbb{E} \left[e^{ju^T X} \right] = \int_{\mathbb{R}^n} e^{ju^T x} f_X(x) dx \quad (\text{D.56})$$

$$\Psi_X(u) = \ln \left(\mathbb{E} \left[e^{ju^T X} \right] \right) \quad (\text{D.57})$$

où u est le vecteur colonne $(u_1, u_2, \dots, u_n)^T$.

Le *moment croisé d'ordre q* de X est constitué des moments d'ordres q_1, q_2, \dots, q_n tels que $q = \sum_i q_i$:

$$\mathbb{E} [X_1^{q_1} X_2^{q_2} \dots X_n^{q_n}] \quad (\text{D.58})$$

Le moment centré d'ordre q est $\mathbb{E} [X_{c,1}^{q_1} X_{c,2}^{q_2} \dots X_{c,n}^{q_n}]$ avec $X_{c,i} = X_i - \mathbb{E}[X_i]$.

De la même manière, le *cumulant croisé d'ordre q* est :

$$c \left[\underbrace{X_1, \dots, X_1}_{q_1 \text{ fois}}, \underbrace{X_2, \dots, X_2}_{q_2 \text{ fois}}, \dots, \underbrace{X_n, \dots, X_n}_{q_n \text{ fois}} \right] \quad (\text{D.59})$$

D.1.4.3 Matrice de covariance et de corrélation

Le moment d'ordre un est l'espérance du vecteur aléatoire X et regroupe l'ensemble des moments d'ordre un de X :

$$\mathbb{E}[X] = (\mathbb{E}[X_1], \mathbb{E}[X_2], \dots, \mathbb{E}[X_n])^T \quad (\text{D.60})$$

La *matrice de corrélation*, ou moment d'ordre deux, est donnée par :

$$m_2[X] = \mathbb{E}[X X^T] = \begin{pmatrix} \mathbb{E}[X_1^2] & \mathbb{E}[X_1 X_2] & \cdots & \mathbb{E}[X_1 X_n] \\ \mathbb{E}[X_2 X_1] & \mathbb{E}[X_2^2] & \ddots & \mathbb{E}[X_2 X_n] \\ \vdots & \ddots & \ddots & \vdots \\ \mathbb{E}[X_n X_1] & \mathbb{E}[X_n X_2] & \cdots & \mathbb{E}[X_n^2] \end{pmatrix} \quad (\text{D.61})$$

Le moment centré d'ordre deux, qui est égal au cumulants d'ordre deux, correspond à la *matrice de covariance* :

$$\begin{aligned} c_2[X] = \mu_2[X] &= \mathbb{E}[X_c X_c^T] = \mathbb{E} \left[(X - \mathbb{E}[X]) (X - \mathbb{E}[X])^T \right] \\ &= \begin{pmatrix} \sigma_{X_1}^2 & \text{Cov}[X_1 X_2] & \cdots & \text{Cov}[X_1 X_n] \\ \text{Cov}[X_2 X_1] & \sigma_{X_2}^2 & \ddots & \text{Cov}[X_2 X_n] \\ \vdots & \ddots & \ddots & \vdots \\ \text{Cov}[X_n X_1] & \text{Cov}[X_n X_2] & \cdots & \sigma_{X_n}^2 \end{pmatrix} \end{aligned} \quad (\text{D.62})$$

Les matrices de covariance et de corrélation sont deux matrices symétriques et définies semi-positives [107], ce qui implique que leur déterminant et leurs valeurs propres sont positifs. Lorsque la matrice de covariance $c_2[X]$ est diagonale, les composantes X_i du vecteur X sont décorréliées (et réciproquement).

De manière similaire, on pourrait exprimer les moments et les cumulants d'ordre q supérieur à deux et qui seront donc des tenseurs d'ordre q . Par exemple, le cumulants d'ordre trois est un tenseur d'ordre trois, c'est-à-dire l'équivalent d'un cube de dimension $n \times n \times n$.

D.1.5 Estimations

L'utilisation des statistiques d'ordre supérieur passe par leur estimation [88, 108, 111]. Cette sous-section définit les notions d'échantillon, de statistique et d'estimateur, puis elle présente quelques éléments sur l'estimation, essentiellement dans le cas scalaire. En pratique, les traitements commencent généralement par une étape de centrage des données à traiter afin de simplifier grandement les développements théoriques. Lorsque le nombre de données est suffisamment grand, on peut supposer que le centrage est effectué de manière parfaite.

D.1.5.1 Echantillon, statistique et estimateur

En pratique, les variables aléatoires sont souvent un ensemble de tirages indépendants selon une même loi inconnue. L'objectif est alors d'estimer le comportement de cette loi. Un *échantillon* (X_1, X_2, \dots, X_n) d'une variable aléatoire X est un vecteur aléatoire réel à n composantes indépendantes et de même loi que X [107].

Une *statistique* U_n est une variable aléatoire définie à partir d'un échantillon (X_1, X_2, \dots, X_n) . Soit $f_{X,\theta}(x)$ la fonction de densité d'une variable aléatoire X dépendant du paramètre inconnu θ . La statistique U_n est un *estimateur* de θ si la valeur observée de U_n fournit une valeur considérée comme approchée de θ .

Un estimateur est dit *cohérent* si la statistique U_n converge en probabilité vers θ quand n tend vers l'infini :

$$U_n \xrightarrow[n \rightarrow +\infty]{} \theta \quad (\text{D.63})$$

L'estimateur est *sans biais* si son espérance mathématique $E[U_n]$ est égale à θ :

$$E[U_n] = \theta \quad (\text{D.64})$$

D.1.5.2 Estimation des moments et des cumulants

Soit X une variable aléatoire scalaire centrée, et (X_1, X_2, \dots, X_n) , n réalisations de X . L'estimateur le plus simple des moments d'ordre q de X est :

$$\widehat{\mu}_q[X] = \frac{1}{n} \sum_{i=1}^n X_i^q \quad (\text{D.65})$$

Cet estimateur est non biaisé puisque $E[\widehat{\mu}_q[X]] = \mu_q[X]$. En utilisant cet estimateur des moments, il est donc possible de construire des estimateurs de cumulants à partir de la formule de Leonov et Shiryayev (équation (D.30)).

Sachant que les cumulants d'ordre inférieur ou égal à trois sont égaux aux moments centrés de même ordre, nous allons nous intéresser directement au cumulants d'ordre quatre qui est lié aux moments centrés d'ordre inférieur ou égal à quatre par l'équation (D.32). Un estimateur simple du cumulants d'ordre quatre est donc :

$$\widehat{c}_4 = \widehat{\mu}_4 - 3\widehat{\mu}_2^2 \quad (\text{D.66})$$

En étudiant son biais, on peut montrer que [88] :

$$E[\widehat{c}_4] = c_4 - 3\frac{c_4 + 2\mu_2^2}{n} \quad (\text{D.67})$$

Cet estimateur est donc biaisé mais le biais tend vers zéro lorsque le nombre d'échantillons n tend vers l'infini. On peut donc considérer, que si le nombre de données est grand, \widehat{c}_4 est un estimateur approximativement non biaisé.

Dans le cas où le nombre d'échantillons est limité, il existe des estimateurs non biaisés des cumulants : les k -statistiques. Les auteurs de [88, 108, 111] donnent et comparent différents estimateurs de cumulants, notamment pour l'ordre quatre. Toutefois, dans cette thèse, nous nous sommes limités à l'estimateur \hat{c}_4 mais l'utilisation des k -statistiques pourrait être une perspective d'amélioration de nos descripteurs basés sur les cumulants.

D.2 Application aux signaux temporels

Dans les sections précédentes, nous avons abordé certaines notions fondamentales de la théorie des probabilités sur laquelle repose une partie de la problématique du traitement du signal. Les variables aléatoires et les grandeurs qui s'y rapportent (fonction de répartition, densité de probabilité, moments, cumulants) ont ainsi été abordées de manière statique, comme si l'on établissait leurs caractéristiques statistiques à un instant donné. En pratique, les phénomènes physiques que nous étudions (la parole par exemple) dépendent également du temps. C'est l'origine de la définition des processus aléatoires que nous allons considérer dans la sous-section D.2.1. Nous y évoquerons leurs principales caractéristiques et leurs moments d'ordre un et deux. Ensuite, en sous-section D.2.2, nous envisagerons les ordres supérieurs à deux plus connus sous le nom de multicorrélations.

D.2.1 Caractérisation des processus aléatoires

Avant de s'intéresser à leur caractérisation, nous allons d'abord définir les notions de signaux aléatoires et de processus aléatoires.

D.2.1.1 Définitions préliminaires

Un signal temporel est dit *déterministe* si, à chaque instant t , sa valeur peut être déterminée de manière certaine par un modèle mathématique [120], comme :

$$s(t) = \begin{cases} A_0 \sin(2\pi f_0 t) + A_1 \sin(2\pi f_1 t) & \text{si } 0 < t < 0.5 \\ A_0 \sin(2\pi f_0 t) + A_2 \sin(2\pi f_1 t) & \text{si } 0.5 < t < 1 \\ 0 & \text{sinon} \end{cases} \quad (\text{D.68})$$

avec trois amplitudes (A_0, A_1, A_2) et trois fréquences (f_0, f_1, f_2) constantes. Ce signal est illustré et analysé (dans les domaines temporel, fréquentiel, temps-fréquence, temps-échelle et des ondelettes) à la figure C.1 de l'annexe C.1.

On parlera de *signal aléatoire* ou *stochastique* si à chaque instant, sa valeur ne peut être déterminée de façon certaine car elle est dépendante du hasard [50]. Cette incertitude provient d'un manque d'information ou de l'impossibilité de décrire un phénomène physique du fait de sa trop grande complexité.

Un *processus aléatoire* (ou stochastique) est une famille de variables indexées par un paramètre t (le temps par exemple) :

$$\begin{aligned} X : \mathbb{R} \times \Omega &\rightarrow \mathbb{R} \\ t \times \omega &\mapsto X(t, \omega) \end{aligned} \quad (\text{D.69})$$

où t représente usuellement l'indice temporel et ω l'indice aléatoire d'expérience [50]. Lorsque t est discret (Ω dénombrable), on parle de *suite aléatoire*.

On appelle *trajectoire* $X_i(t)$ associée à l'expérience ω_i la réalisation du processus $X(t, \omega_i)$:

$$\begin{aligned} X_i : \mathbb{R} &\rightarrow \mathbb{R} \\ t &\mapsto X(t, \omega_i) \end{aligned} \quad (\text{D.70})$$

En pratique, une trajectoire du processus $X(t, \omega)$ est considérée comme un signal déterministe bien qu'il n'existe pas forcément de formulation analytique.

De manière similaire, on peut définir la variable aléatoire X_t , appelée *t-section*, pour un instant t par :

$$\begin{aligned} X_t : \Omega &\rightarrow \mathbb{R} \\ \omega &\mapsto X(t, \omega) \end{aligned} \quad (\text{D.71})$$

D.2.1.2 Propriétés statistiques des deux premiers moments

Les propriétés statistiques d'ordre un d'un processus aléatoire continu (ou discret) sont déterminées en étudiant le processus pour un temps t fixé [120]. Le processus aléatoire devient alors une variable aléatoire et toutes les définitions et propriétés vues à la sous-section précédente s'appliquent. La fonction de répartition d'un processus aléatoire est donc la même que celle d'une variable aléatoire :

$$F_X(x, t) = P(X(t) \leq x) \quad (\text{D.72})$$

Cela correspond à la probabilité que $X(t, \omega)$ se trouve en dessous d'un seuil donné pour un temps t donné.

De la même manière, la densité de probabilité correspond à la probabilité qu'a le processus aléatoire de passer dans une fenêtre dx donnée à un temps t :

$$f_X(x, t) = \frac{dF_X(x, t)}{dx} \quad (\text{D.73})$$

Par définition, la valeur moyenne d'un processus aléatoire $X(t)$ pris au temps t est :

$$E[X(t)] = \int_{-\infty}^{+\infty} x f_X(x, t) dx \quad (\text{D.74})$$

Les statistiques d'ordre deux d'un processus aléatoire sont déterminées en étudiant le processus à deux temps différents. Suivant le même principe que l'ordre un, la fonction de répartition d'ordre deux représente la probabilité que le processus passe en dessous de deux seuils à deux instants différents :

$$F_X(x_1, x_2, t_1, t_2) = P(X(t_1) \leq x_1, X(t_2) \leq x_2) \quad (\text{D.75})$$

De même la densité de probabilité correspond à la probabilité qu'a le processus de passer dans deux fenêtres à deux temps différents :

$$f_X(x_1, x_2, t_1, t_2) = \frac{d^2 F_X(x_1, x_2, t_1, t_2)}{dx_1 dx_2} \quad (\text{D.76})$$

Par définition, le moment d'ordre deux est :

$$m_2[X](t) = E[X(t)^2] = \int_{-\infty}^{+\infty} x^2 f_X(x, t) dx \quad (\text{D.77})$$

Considérons deux échantillons $X(t_1)$ et $X(t_2)$ d'un processus aléatoire pris aux instants t_1, t_2 , la fonction d'autocorrélation permet de mesurer le degré de dépendance (ou d'énergie mutuelle) qu'il existe entre ces deux échantillons [120]. Elle est définie par :

$$M_2[X](t_1, t_2) = E[X(t_1) X(t_2)] \quad (\text{D.78})$$

$$= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x_1 x_2 f_X(x_1, x_2, t_1, t_2) dx_1 dx_2 \quad (\text{D.79})$$

Il est possible de généraliser la notion d'autocorrélation aux processus aléatoires complexes par la relation :

$$M_2[X](t_1, t_2) = E[X(t_1) X^*(t_2)] \quad (\text{D.80})$$

où l'astérisque en exposant indique qu'il s'agit du conjugué de X_i . Comme nos images sont à valeurs réelles, la suite de cette section ne considère que des signaux aléatoires réels (d'où $X_i^* = X_i$).

De même, l'extension aux vecteurs aléatoires de dimension n permet de définir la matrice d'autocorrélation $(M_2[X])_{0 < i, j \leq n}$:

$$M_2[X](t_i, t_j) = E \left[X(t_i), X^T(t_j) \right] \quad (\text{D.81})$$

En centrant les variables aléatoires par rapport à leur valeur moyenne, nous obtenons alors la fonction de covariance (ou matrice de covariance pour un vecteur aléatoire).

D.2.2 Ergodicité et stationnarité

En pratique, l'étude d'un signal se résume généralement à l'exploitation de son information temporelle, le plus souvent tronquée (*e.g.* pour $t \in [0, T]$) [50]. Afin d'apporter des informations sur les grandeurs statistiques d'un processus aléatoire, il faudrait disposer de plusieurs réalisations d'un même processus à un temps donné. Or ce n'est pas souvent réalisable car nous avons seulement accès à une réalisation particulière du processus. Dans cette sous-section, nous allons aborder deux notions importantes des processus aléatoires qui permettent de simplifier un problème : l'ergodicité et la stationnarité.

D.2.2.1 Ergodicité

L'*ergodicité* est la propriété d'un système qui tend vers un état limite indépendant de la situation initiale (exemple du battage d'un jeu de carte). Autrement dit, le caractère ergodique d'un signal aléatoire consiste à admettre que son évolution au cours du temps apporte une information du même type que celle obtenue à un instant donné d'un ensemble de réalisations de la variable aléatoire [50]. La propriété d'ergodicité permet de généraliser les moments temporels calculés sur une trajectoire à toutes les trajectoires possibles du processus.

Un processus aléatoire $X(t, \omega_i)$ est dit *ergodique* si tous les moments temporels existent et sont indépendants de l'expérience ω_i considérée (sauf pour un ensemble de trajectoires de probabilité nulle) .

Le processus aléatoire $X(t, \omega_i)$ est dit *ergodique à l'ordre un* si sa moyenne temporelle est indépendante de l'expérience ω_i considérée. L'espérance mathématique d'un signal $X_i(t)$ ergodique (au moins à l'ordre un) peut être approchée par sa moyenne temporelle $M_1[X_i]$:

$$m_1[X_i] = E[X_i] \approx E[X_i(t)] = M_1[X_i] \quad (\text{D.82})$$

Un processus aléatoire $X(t, \omega_i)$ est dit *ergodique à l'ordre deux* s'il est ergodique à l'ordre 1 et que son autocorrélation (temporelle) est indépendante de l'expérience ω_i .

D'une manière générale, un processus est dit *ergodique à l'ordre n* si ses moments statistiques jusqu'à l'ordre n coïncident asymptotiquement avec ses moments temporels d'ordre n .

D.2.2.2 Stationnarité

Un processus aléatoire est *stationnaire* lorsque ses statistiques sont invariantes dans le temps et, plus précisément, indépendantes de l'origine des temps. La propriété de stationnarité est donc à la t -section ce qu'est l'ergodicité à l'expérience aléatoire : elle permet d'établir les caractéristiques statistiques d'un processus aléatoire à partir d'une seule t -section.

Un processus aléatoire $X(t, \omega_i)$ est dit *stationnaire au sens strict* si ses caractéristiques probabilistes sont invariantes pour tout changement de l'origine des temps t_0 :

$$F_{X_{t_1}, X_{t_2}, \dots, X_{t_n}}(x_{t_1}, x_{t_2}, \dots, x_{t_n}) = F_{X_{t_1+t_0}, X_{t_2+t_0}, \dots, X_{t_n+t_0}}(x_{t_1}, x_{t_2}, \dots, x_{t_n}) \quad (\text{D.83})$$

pour tout indice i (entre 1 et n) et où $F_{X_{t_1}, \dots, X_{t_n}}(x_{t_1}, \dots, x_{t_n})$ est la fonction de répartition conjointe du vecteur aléatoire $X = (X_{t_1}, \dots, X_{t_n})^T$.

De la même manière que pour l'ergodicité, un processus $X(t, \omega_i)$ est dit *stationnaire à l'ordre un* si son espérance mathématique ne dépend pas du temps. La moyenne temporelle d'un signal $X_i(t)$ ergodique (au moins à l'ordre un) peut donc être approchée par son espérance mathématique $m_1[X_i]$.

Un processus aléatoire $X(t, \omega_i)$ est dit *stationnaire à l'ordre deux*, s'il est stationnaire à l'ordre un et si la corrélation (ou moment croisé d'ordre deux) entre deux t - et $(t + \tau)$ -sections ne dépend que de la différence τ .

De manière générale, un processus aléatoire est *stationnaire jusqu'à l'ordre n* si ses moments statistiques sont indépendants de l'origine des temps jusqu'à l'ordre n . Ceci implique que le moment d'ordre k d'un processus stationnaire jusqu'à cet ordre ne dépend plus que de $k - 1$ variables.

D.2.2.3 Ergodicité et stationnarité

Les deux notions de stationnarité et d'ergodicité sont indépendantes. Toutefois, si un processus aléatoire $X(t, \omega_i)$ est stationnaire et ergodique à l'ordre deux (ou plus), alors ses moments statistiques et temporels d'ordre un et deux sont presque sûrement égaux [50]. Par conséquent, les propriétés statistiques d'un processus aléatoire stationnaire et ergodique peuvent être obtenues à partir de l'observation d'une seule et unique trajectoire sur une seule et unique t -section.

D.2.3 Moments temporels d'ordres un et deux d'un signal stationnaire

Dans le cas où le processus aléatoire est stationnaire, nous pouvons réécrire les relations précédentes (cf. sous-section D.2.1) sans la variable de temps :

$$\begin{aligned} f_X(x, t) &= f_X(x) \\ f_X(x_1, x_2, t_1, t_2) &= f_X(x_1, x_2, \tau) \end{aligned}$$

où $f_X(x_1, x_2, \tau)$ ne dépend plus de l'origine des temps mais de l'écart (ou retard) $\tau = t_2 - t_1$.

De même, on peut définir la *moyenne temporelle* (ou moment temporel d'ordre un) de la trajectoire $X_i(t)$ par :

$$M_1[X_i] = \text{E}[X_i(t)] \quad (\text{D.84})$$

Le moment temporel d'ordre deux, l'*autocorrélation* de la trajectoire $X_i(t)$ est définie par :

$$M_2[X_i](\tau) = \text{E}[X_i^*(t)X_i(t + \tau)] \quad (\text{D.85})$$

pour tout retard $\tau \in \mathbb{R}$. L'astérisque en exposant indique que c'est le conjugué de X_i qu'il faut prendre. Comme on ne considère que des signaux aléatoires réels, on a $X_i^* = X_i$. Par la suite, le conjugué ne sera plus indiqué dans les définitions, bien qu'il faudrait le prendre en compte dans le cas de signaux complexes.

Pour un signal à puissance finie, les moments temporels d'ordres un et deux sont définis par :

$$\begin{aligned} M_1[X_i] &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} X_i(t) dt \\ M_2[X_i](\tau) &= \lim_{T \rightarrow \infty} \left(\frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} X_i(t + \tau) X_i^*(t) dt \right) \end{aligned}$$

La figure D.3 illustre le principe de l'autocorrélation avec pour signal $s(t)$ l'ondelette du chapeau mexicain définie par l'équation (C.7). La figure D.3.a représente l'ondelette dans le domaine temporelle avec ses répliques retardées. La figure D.3.b donne l'autocorrélation en fonction du retard de la réplique.

Tout comme on a défini le moment croisé d'ordre deux entre un couple de variables aléatoires (covariance en section D.1.3.2 équation (D.45)), il est possible de calculer la corrélation croisée entre deux trajectoires (ou deux signaux temporels) $X(t)$ et $Y(t)$ stationnaires (définition en section D.2.2.2) donnée par :

$$M_2[X, Y](\tau) = E [X^*(t)Y(t + \tau)] \quad (D.86)$$

La corrélation croisée permet d'estimer la ressemblance entre les deux trajectoires pour chaque retard. Elle peut par exemple être utilisée pour recaler l'origine d'un signal par rapport à une référence.

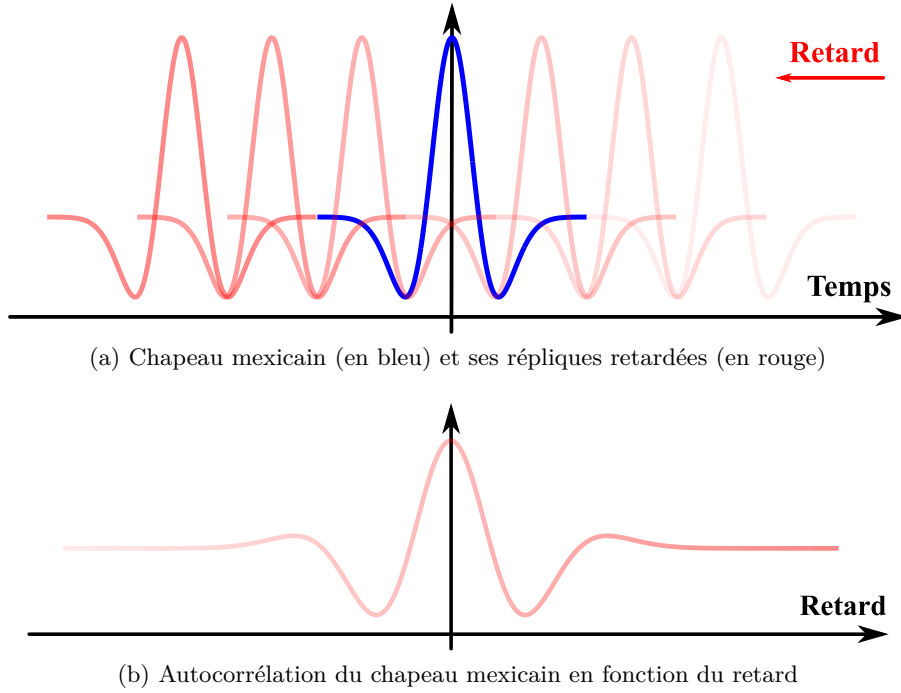


FIGURE D.3 – Principe de l'autocorrélation : chaque retard τ de l'autocorrélation correspond à l'espérance du produit du signal avec sa réplique retardée de τ sur l'ensemble de l'espace temporel.

On retrouve également le concept de cumulants temporels d'ordre deux, qui, pour une trajectoire $X_i(t)$, est défini par :

$$\mathcal{C}_2[X_i](\tau) = \mathcal{C} [X_i(t), X_i(t + \tau)] = M_2[X_i](\tau) - M_1^2[X_i] = \mathcal{M}_2[X_i](\tau) \quad (D.87)$$

où $\mathcal{M}_2[X_i](\tau)$ désigne le moment centré d'ordre deux :

$$\mathcal{M}_2[X_i](\tau) = E [(X_i^*(t) - M_1[X_i]) (X_i(t + \tau) - M_1[X_i])] \quad (D.88)$$

De même, le cumulants croisés d'ordre deux de deux trajectoires $X(t)$ et $Y(t)$ est :

$$\mathcal{C}_2[X, Y](\tau) = \mathcal{C} [X(t), Y(t + \tau)] = M_2[X, Y](\tau) - M_1[X]M_1[Y] = \mathcal{M}_2[X, Y](\tau) \quad (D.89)$$

Un processus aléatoire $X(t, \omega)$ ne sera entièrement caractérisé que si l'on est capable d'en donner une description statistique complète, autrement dit, si, pour tout k -uplets d'instant (t_1, t_2, \dots, t_k) , on connaît la densité de probabilité, la fonction de répartition ou la fonction caractéristique du vecteur aléatoire $X_t = [X_{t_1}, X_{t_2}, \dots, X_{t_k}]^T$. On se retrouve donc dans le cas de la section D.1.4 où toutes les définitions et propriétés évoquées sur les vecteurs aléatoires s'appliquent sur X_t : moments, cumulants, etc.

D.2.4 Multicorrélations

Dans cette sous-section, nous allons nous consacrer uniquement aux moments et cumulants temporels, $X(t, \omega)$ sera noté $X(t)$ pour simplifier les équations et les signaux temporels seront centrés (c'est-à-dire de moyenne nulle).

A l'instar de la fonction de corrélation d'ordre deux, on peut définir des fonctions de multicorrélation d'ordre supérieur. Aux ordres deux et trois, les moments centrés et les cumulants sont confondus, de sorte qu'il n'existe pas d'ambiguïté. En revanche aux ordres plus grands, il faudra distinguer les multicorrélations tirées des cumulants ou des moments centrés.

D.2.4.1 Temps continu

Soit un signal aléatoire $X(t)$ à valeurs réelles, la *multicorrélation cumulante* d'ordre q de $X(t)$ est le cumulante d'ordre q des valeurs (variables aléatoires) du signal $X(t)$ aux instants t_0, t_1, \dots, t_{q-1} [88] :

$$\mathcal{C}_q[X](t) = \mathcal{C}[X(t_0), X(t_1), \dots, X(t_{q-1})] \quad (\text{D.90})$$

où $t = (t_0, t_1, \dots, t_{q-1})$. Dans le cas général, la multicorrélation d'ordre q est fonction d'un temps absolu t_0 et des $q - 1$ temps t_i .

Dans le cas particulier où le signal $X(t)$ est stationnaire au sens strict, ses statistiques sont donc invariantes par changement de l'origine des temps. En choisissant un temps t_0 comme origine, la multicorrélation n'est alors fonction que des $q - 1$ écarts de temps $\tau_i = t - t_i$, appelés retards. La multicorrélation d'ordre q est alors donnée par :

$$\mathcal{C}_q[X](\tau) = \mathcal{C}[X(t), X(t + \tau_1), \dots, X(t + \tau_{q-1})] \quad (\text{D.91})$$

où $\tau = (\tau_1, \tau_2, \dots, \tau_{q-1})$. Le moment temporel d'ordre q est défini par :

$$\mathcal{M}_q[X](\tau) = \text{E}[X(t)X(t + \tau_1) \dots X(t + \tau_{q-1})] \quad (\text{D.92})$$

On retrouve la corrélation classique à l'ordre deux définie en section D.2.3, mais aussi la *bicorrélation* (cumulant d'ordre trois)

$$\mathcal{C}_3[X](\tau_1, \tau_2) = \mathcal{C}[X(t), X(t + \tau_1), X(t + \tau_2)] \quad (\text{D.93})$$

et la *tricorrélation* (cumulant d'ordre quatre)

$$\mathcal{C}_4[X](\tau_1, \tau_2, \tau_3) = \mathcal{C}[X(t), X(t + \tau_1), X(t + \tau_2), X(t + \tau_3)] \quad (\text{D.94})$$

Les multicorrélations, définies à partir des cumulants, peuvent s'exprimer en fonction des moments en utilisant les relations entre les cumulants et les moments données à la section D.1.2. Lorsque les signaux sont centrés, la bicorrélation et la tricorrélation (et les ordres supérieurs) s'expriment en fonction des moments par :

$$\mathcal{C}_3[X](\tau_1, \tau_2) = \text{E}[X(t)X(t + \tau_1)X(t + \tau_2)] = \mathcal{M}_3[X](\tau_1, \tau_2) \quad (\text{D.95})$$

$$\mathcal{C}_4[X](\tau_1, \tau_2, \tau_3) = \mathcal{M}_4[X](\tau_1, \tau_2, \tau_3) \quad (\text{D.96})$$

$$- \mathcal{M}_2[X](\tau_1)\mathcal{M}_2[X](\tau_2 - \tau_3) \quad (\text{D.97})$$

$$- \mathcal{M}_2[X](\tau_2)\mathcal{M}_2[X](\tau_1 - \tau_3) \quad (\text{D.98})$$

$$- \mathcal{M}_2[X](\tau_3)\mathcal{M}_2[X](\tau_1 - \tau_2) \quad (\text{D.99})$$

D.2.4.2 Temps discret

Supposons que nous avons un signal stationnaire à temps continu $X(t)$ que nous échantillons pour donner un signal à temps discret (numérique) $X(n)$. Pour simplifier les notations, la période d'échantillonnage T_E est sous-entendue et l'instant nT_E est noté n . La multicorrélation cumulée du signal discret $X(n)$ est donnée par [88] :

$$\mathcal{C}_q[X](i) = \mathcal{C}[X(n), X(n+i_1), \dots, X(n+i_{q-1})] \quad (\text{D.100})$$

où $i = (i_1, i_2, \dots, i_{q-1})$. Son moment temporel d'ordre q est :

$$\mathcal{M}_q[X](\tau) = \text{E}[X(n)X(n+i_1)\dots X(n+i_{q-1})] \quad (\text{D.101})$$

D.2.4.3 Estimation

De la même manière que pour les moments et cumulants statistiques, l'utilisation d'un estimateur est souvent indispensable pour approcher les multicorrélations. Pour ce faire, nous proposons d'appliquer le même estimateur que celui exposé en section D.1.5. Le moment temporel centré d'ordre q est estimé par :

$$\widehat{\mathcal{M}}_q[X](i) = \frac{1}{N} \sum_{n \in D(N,i)} X(n)X(n+i_1)\dots X(n+i_{q-1}) \quad (\text{D.102})$$

où N est le nombre de points du signal observé et $D(N, i)$ est le domaine des valeurs parcourus par n , qui dépend donc du nombre de points N et des valeurs des retards $i = (i_1, i_2, \dots, i_{q-1})$ pour lesquels le moment est calculé.

Par exemple, la bicorrélation peut donc être estimée par :

$$\widehat{\mathcal{C}}_3[X](i_1, i_2) = \widehat{\mathcal{M}}_3[X](i_1, i_2) = \frac{1}{N} \sum_{n \in D(N, i_1, i_2)} X(n)X(n+i_1)X(n+i_2) \quad (\text{D.103})$$

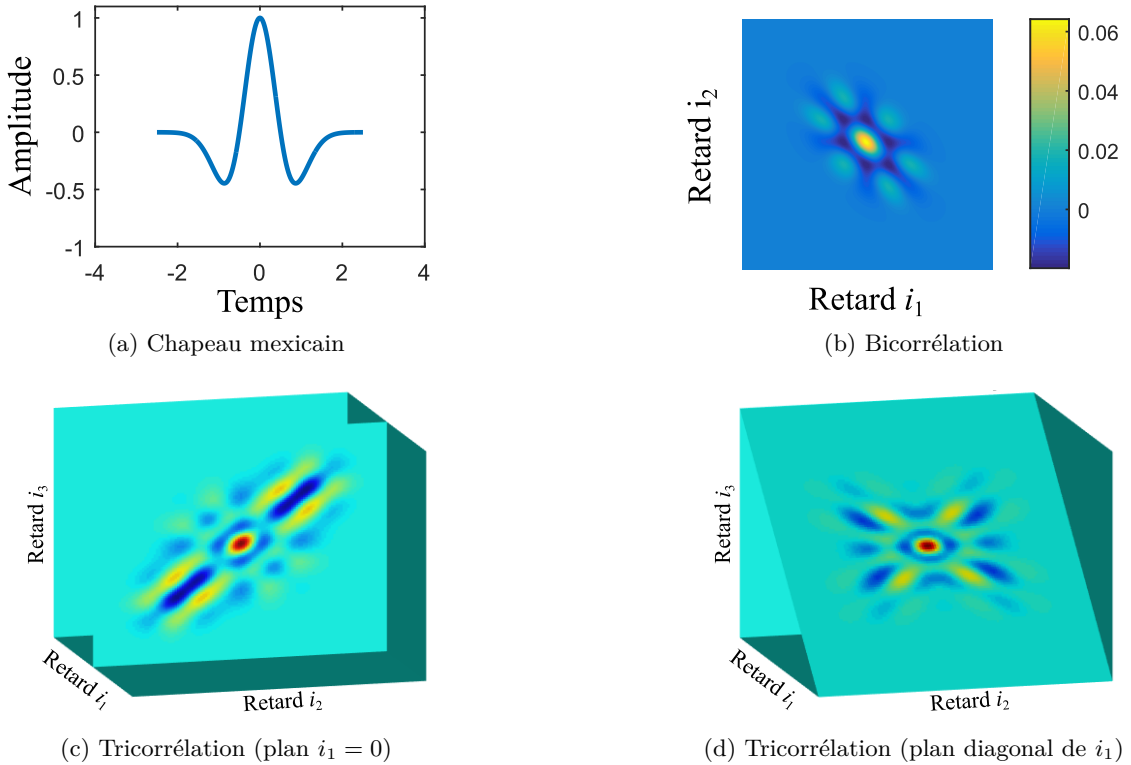


FIGURE D.4 – Bicorrélation et tricorrélation du chapeau mexicain.

La figure D.4.b affiche la bicorrélation du chapeau mexicain (figure D.4.a) dont l'autocorrélation était illustrée sur la figure D.3. La bicorrélation est donc une image dont les abscisses et les ordonnées sont les retards i_1 et i_2 . La tricorrélation, qui dépend de trois retards (i_1, i_2, i_3) , est un tenseur d'ordre trois (*i.e.* l'équivalent d'un cube), dont deux plans de coupe sont représentés sur les figures D.4.c et D.4.d. Les multicorrélations d'ordre supérieur q sont donc des tenseurs d'ordre $q - 1$ difficilement représentables dans l'espace (seulement par plan de coupe) et contenant un nombre de plus en plus important de valeurs.

Conclusion

La première partie de cet annexe est une introduction aux statistiques d'ordre supérieur qui traite essentiellement des principaux outils utilisés dans les techniques d'ordre supérieur : les moments et les cumulants. Cette présentation rapide est axée sur la définition des notions mathématiques utilisées dans la suite de cette section et de leurs propriétés. La seconde partie de cette section est consacrée aux signaux à une dimension et à leur caractérisation par les outils statistiques vue en première partie. Les propriétés d'ordre supérieur des signaux aléatoires peuvent en effet être décrites, comme le sont les propriétés d'ordre deux, dans le domaine du temps : on parle alors de multicorrélations. Il est à noter qu'il est également possible d'étudier ces propriétés dans le domaine des fréquences, les multispectres, ce qui pourrait constituer une perspective intéressante pour la suite des travaux de notre projet. La suite de cet annexe fait l'objet de la section 4.2 et consiste à étendre les définitions exposées pour les signaux temporels aux signaux bidimensionnels : les images.

Annexe E

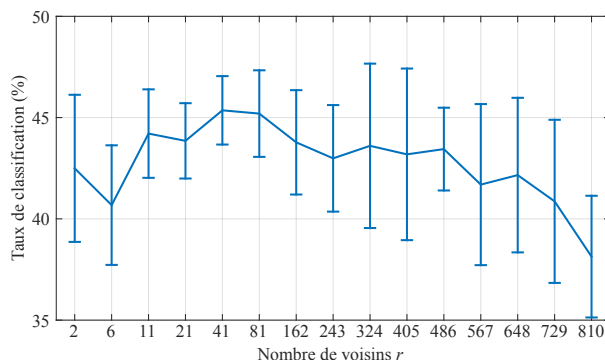
Choix du paramètre de dispersion en classification spectrale

En classification spectrale, il est nécessaire de choisir une distance et une fonction de similarité. Dans notre étude (*cf.* section 2.4.2.3), nous avons opté pour la similitude gaussienne f_g couplée avec la distance euclidienne d_{euc} qui, pour deux images I et J de taille N , sont définies par :

$$f_g(I, J) = \exp\left(-\frac{d_{euc}^2(I, J)}{2\sigma^2}\right) \text{ avec } d_{euc}(I, J) = \sqrt{\sum_{i=1}^N (I_i - J_i)^2} \quad (\text{E.1})$$

Le coefficient σ est un paramètre qui traduit la dispersion des données. Pour définir ce paramètre, une solution consiste à définir σ^2 comme le produit $\sigma_I \sigma_J$, où ces deux coefficients dépendent du voisinage de chaque point [165]. Le coefficient σ_I est alors déterminé en calculant la distance euclidienne par $\sigma_I = d_{euc}(I, I_r)$ où I_r correspond au r -ième voisin le plus proche de I .

La figure E.1.a montre l'évolution du taux de bonne classification en fonction du rang r sur la base KTH-TIPS. Chaque classe comportant 81 images, le fait de fixer r au nombre d'images par classe semble être le choix le plus logique. Les résultats de la figure E.1.a confirment que ce choix est cohérent avec un taux de classification d'environ 45 %. On constate tout de même que les performances du *Scattering Operator* (SO) en supervisé (avec apprentissage par SMA) sont nettement supérieures à la classification spectrale. En observant la matrice de confusion (E.1.b) de la classification spectrale, on constate que certaines classes (#4 Coton, #8 Papier de verre, #9 Eponge et #10 Polystyrène) sont très mal classées ce qui fait chuter le taux de classification.



(a) Moyenne et écart type du taux de classification en fonction du paramètre r sur 10 tests

S1	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10
#1	90	0	0	0	8	3	0	0	0	0
#2	0	65	7	2	10	6	0	4	2	3
#3	0	11	49	0	18	7	0	7	7	0
#4	0	1	1	18	14	17	0	14	15	21
#5	7	3	1	0	54	29	0	3	4	0
#6	0	1	0	8	22	39	19	0	1	8
#7	0	0	2	11	0	12	50	13	0	12
#8	0	1	10	17	2	4	0	27	19	20
#9	0	0	22	13	17	5	0	4	27	12
#10	0	0	3	21	7	13	0	15	10	31

(b) Matrice de confusion pour $r = 81$

FIGURE E.1 – Influence du nombre de voisins r considérés dans le calcul du paramètre de dispersion en classification spectrale : les deux meilleurs taux de classification sont obtenus pour $r = 41$ (45.36 ± 1.17 %) et $r = 81$ (45.20 ± 2.14 %). Une seconde simulation comprenant 10 tests a permis d'établir cette matrice de confusion avec $r = 81$ et a donné un taux de classification équivalent de 44.89 ± 2.20 %.

Annexe F

Simulations sur une base réduite de thrombi

Glossaire de l'annexe B

CS Classification Spectrale (*spectral clustering* en anglais)

EP Embolie Pulmonaire

KTH-TIPS Base de textures créée par des membres de l'Institut royal de technologie de Stockholm (*Kungliga Tekniska Högskolan*) et contenant dix classes de matériaux observés sous différentes luminosités, orientations et échelles (*Textured under varying Illumination, Pose and Scale*) [56].

SMA Classification par Sélection de Modèles Affines

SO *Scattering Operator* : réseau convolutif de transformée en ondelettes et d'opérateurs contractants

TCD Transformée en Cosinus Discrète

TVP Thrombose Veineuse Profonde

Introduction

Dans cette annexe, nous souhaitons classifier, de façon automatique, les images échographiques et élastographiques en fonction de l'origine du thrombus et la survenue d'embolie pulmonaire des patients. Ces images sont analysées par le *scattering operator* (décrit à la section 4.1) et classées par la classification spectrale (*cf.* section 2.4.2.3). Ces simulations ont été réalisées lors de la deuxième année et font l'objet de l'article de conférence [12]. La base de données utilisée est une base réduite de la base acquise avec le protocole 2B (*cf.* section 1.3.2.2). En effet, pour ces simulations, nous avons sélectionné les données collectées avec les mêmes paramètres d'acquisition afin de s'affranchir du biais que pourraient amener des changements de paramètres. La contre-partie est que cette base de test est de taille restreinte mais elle reste plus grande que celle du protocole 2A et possède des élastographies de meilleure qualité que le protocole 1. Cet annexe s'articule en quatre sections. La section F.1 décrit d'abord le protocole expérimental (la base et le paramétrage des algorithmes). Ensuite, les sections F.2 et F.3 donnent les résultats de classification sur, respectivement, les échographies et les élastographies. Enfin, en section F.4, les coefficients du *scattering operator* de chaque échographie et de chaque élastographie sont fusionnés pour ne former qu'un seul vecteur de descripteurs, puis l'ensemble des vecteurs ainsi formé est classé en deux, trois ou quatre sous-ensembles.

F.1 Protocole expérimental

La base d’images texturées, présentée en section 5.1, a permis d’évaluer le *Scattering Operator* (SO) et la Classification Spectrale (CS) afin de sélectionner les paramètres les plus performants (section 5.1.2 et annexe E). Ces études ont montré que le SO est beaucoup plus performant (sur la base KTH-TIPS) s’il est suivi d’un classifieur linéaire supervisé (en l’occurrence la SMA) que d’un algorithme de partitionnement non supervisé comme les k -moyennes. Toutefois, la base de thrombi, considérée dans cette annexe, est trop limitée en nombre d’images par classe pour pouvoir lui appliquer une approche supervisée. Dans ces simulations, nous avons donc opté pour le même jeu de paramètres que celui employé sur la base KTH-TIPS, à savoir huit orientations, cinq échelles et deux ordres de décomposition. La taille des images est fixée de telle sorte à être la plus grande possible, à savoir 64×64 pixels (images redimensionnées par TCD).

Par ailleurs, si les invariances en rotation et en échelle ne semblent pas forcément améliorer les taux de classification (notamment en non supervisé), l’usage du logarithme paraît, lui, beaucoup plus pertinent. Il permet de renforcer l’information de texture présente dans les coefficients d’ordres supérieurs. Or, en échographie, le gain pouvant varier d’un patient à l’autre, il est donc important que l’information de texture prime sur celles de contraste et de luminosité. Dans le même but, les coefficients d’approximation des échographies (information basse-fréquence) sont exclus des vecteurs descripteurs des images. En élastographie, la valeur moyenne des vitesses des ondes de cisaillement est une information beaucoup plus fiable et significative puisqu’elle indique la dureté globale du thrombus. Les coefficients d’approximation sont donc conservés. Si l’utilisation du logarithme s’avère évidente en échographie, la question de son intérêt peut se poser en élastographie. Toutefois, dans le but de caractériser la structure des thrombi, l’emploi du logarithme a été privilégié pour les simulations présentées dans cette annexe (afin de ne pas surcharger avec ceux obtenus également sans logarithme dont les résultats ne sont pas forcément plus significatifs).

Epidémiologie									Echographies			Elasto.	
Classes	ID	Age	Sexe	TVP	Déclencheur	Gén.	Tab.	Mois	Gain	T	L	T	L
Cancer	012	56	M	FCD	Chirurgie	Fam.	Oui	0	73	1	3	1	3
	013	81	F	FCG				0	88	1	4	1	4
	021	61	M	FCG				0	93	2	3	2	3
Ca. EP	011	93	F	FCD				0	80	1		1	
	018	95	F	FCG				0	79	2	1	2	1
Immob.	008	79	F	FCG	Allitement			0	77	1		1	
Im. EP	019	51	M	PD	Plâtre			0	78	1	6	1	6
Gé. EP	020	91	F	FCD		Ant.		0	73	1	4	1	4
	003	37	M	ICG		FLV	Oui	1	87	2	4	2	4
Idiop.	001	88	F	PD				0	73	2	4	2	4
Id. EP	010	57	M	FCG				0	84	1		1	

Tableau F.1 – Liste des acquisitions par patient obtenu avec le protocole 2B : classes, identifiant (ID), âge, sexe (féminin/masculin), TVP (Commune, Droite, Fémorale, Gauche, Iliaque et/ou Poplitée), facteur déclencheur, prédispositions génétiques (antécédents Familiers, antécédents personnel ou Facteur V Leiden), tabac (Tab.), mois d’examen (M), gain, nombre d’échographies et d’élastographies transversales (T) et longitudinales (L).

En ce qui concerne l’algorithme de partitionnement, la classification spectrale surpassant les k -moyennes sur la base d’images texturées, elle est donc appliquée soit directement sur les images elles-mêmes, soit sur les transformées du SO. En accord avec les résultats de l’annexe E, le paramètre de dispersion r a été défini comme le rapport entre le nombre d’images et le nombre de classes. Les images sont partitionnées en deux, trois ou quatre classes. Au-delà de quatre classes, les résultats deviennent très variables en fonction de l’initialisation des k -moyennes (dernière étape de la classification spectrale). Enfin, la base de thrombi considérée est celle obtenue avec le protocole 2B pour des raisons de qualité

des élastographies (par rapport au protocole 1) et du nombre d'images (par rapport au protocole 2A). La liste des patients est donnée dans le tableau F.1. On constate qu'il y a en moyenne un ou deux patients par facteur déclencheur (hormis le cas du cancer) d'où le choix de la classification non supervisée.

F.2 Echogénéicité du thrombus

Dans cette section, nous nous intéressons à l'échogénéicité des thrombi : la sous-section F.2.1 envisage les échographies transversales et la sous-section F.2.2 les échographies longitudinales. Les deux types de coupe sont donc étudiés ici de manière séparée. Néanmoins, des simulations analogues ont été effectuées en mélangeant les échographies transversales et longitudinales sans pour autant obtenir des résultats plus parlants. Par ailleurs, les résultats de ces simulations sont disponibles dans l'article de la conférence ICBSIA (*International Conference on Biomedical Signal and Image Analysis*) [12].

F.2.1 Echographie transversale

Ces simulations visent à identifier des corrélations entre le classement des images de la base et le contexte de la Thrombose Veineuse Profonde (TVP) du patient. Le tableau F.2.a propose une classification « idéale » des échographies transversales en fonction du facteur déclencheur et de la survenue d'une Embolie Pulmonaire (EP). Les échographies sont prétraitées pour limiter l'influence du gain, propre à chaque patient, en égalisant leur histogramme localement avec la distribution de Rayleigh (voir section 3.2.3). Ce classement est indicatif et ne peut être considéré comme une vérité absolue au vue de la complexité du problème énoncé précédemment. Les constatations émises par la suite sont donc à relativiser par rapport aux interrogations liées à la définition des classes et au peu de données comparées.

Classement selon l'origine de la thrombose veineuse profonde en mode transversal															
Origine	Cancer			Cancer et EP			Immob.	Im. EP	Génétique et EP			Idiopathique	Id. EP		
Patient	P12	P13	P21	P11	P18			P08	P19	P20	P03		P01	P10	
Age	56	81	61	93	95			79	51	91	37		88	57	
Gain	73	88	93	93	79			77	78	73	87		73	84	
Image	001	011	020	024	016	019	020	001	019	029	001	005	008	010	005

(a) Classes idéales

Classification spectrale appliquée aux échographies à la taille 64x64 (dctpad) en mode transversal															
2 Classes	1						2								
3 Classes	1				2				3				4		
4 Classes	1		2		3		4		5		6		7		
Origine	Gé. EP	Cancer	Cancer et EP		Cancer	Génétique et EP		Idiopathique	Im. EP	Cancer	Immob.	Id. EP			
Patient	P20	P12	P18		P11	P21		P03	P01	P19	P13	P08	P10		
Age	91	56	95		93	61		37	88	51	81	79	57		
Gain	73	73	79		93	93		87	73	78	88	77	84		
Image	029	001	020	019	016	020	024	001	005	008	010	019	011	001	005

(b) Classes obtenues avec la classification spectrale sur les images elles-mêmes

Classification spectrale appliquée aux scattering coefficients des échographies à la taille 64x64 (dctpad) en mode transversal														
2 Classes	1						2							
3 Classes	1				2				3				4	
4 Classes	1		2		3		4		5		6		7	
Origine	Cancer et EP		Génét. EP	Cancer		Idiopathique		Ca. EP	Immob.	Im. EP	Cancer		Génétique et EP	Id. EP
Patient	P18		P20	P12	P13	P01		P11	P08	P19	P21		P03	P10
Age	95		91	56	81	88		93	79	51	61		37	57
Gain	79		73	73	88	73		93	77	78	93		87	84
Image	020	019	029	001	011	008	010	016	001	019	020	024	001	005

(c) Classes obtenues avec la classification spectrale sur les coefficients du scattering operator

Tableau F.2 – Classification des échographies transversales en 2, 3 et 4 sous-ensembles.

Le tableau F.2.b présente les résultats de la classification spectrale appliquée directement sur les images de thrombi. Les images sont partitionnées en deux, trois et quatre classes. Globalement, nous pouvons voir que les thrombi les plus échogènes sont regroupés ensemble. Les patients P19, P01, P08 et P10 sont très échogènes et sont plus ou moins classés ensemble. Cela pourrait indiquer que les TVP idiopathiques des patients P01 et P10 sont assez anciennes ou dues à une stase veineuse. Le patient P13 semble faire figure d'anomalie parmi ce groupe de patients tant par le contexte de survenue (cancer) que par l'échographie elle-même (moins échogène). Hormis ce patient P13, les TVP dues à un cancer sont plutôt regroupées. Le patient P03 qui présente une mutation génétique favorisant la coagulation du sang est, lui, classée de façon indécise. En revanche, la présence ou non d'EP ne semble pas avoir d'impact sur l'échostructure du thrombus.

Ensuite, le tableau F.2.c donne les classements en deux, trois et quatre classes des descripteurs du SO par CS. Les coefficients d'approximation du SO étant exclus des vecteurs descripteurs, nous obtenons logiquement une classification beaucoup plus basée sur la texture échogène des thrombi. Avec ces nombreux petits motifs, les thrombi des patients P21, P03 et P10 sont toujours classés ensemble. Les imageries avec des structures échogènes plus volumineuse (P01, P11, P08 et P19) forment leur propre classe (en quatre classes) ou sont associées à la classe des motifs plus petits. Les imageries avec le moins de structure échogène sont regroupées dans une autre classe (P18, P20, P12 et P13). Enfin, la première imagerie du patient 18 est isolée des autres imageries car elle est totalement inéchogène. Si le classement des imageries peut se comprendre par rapport à l'échogénicité des thrombi, le lien avec l'épidémiologie des patients reste beaucoup plus compliqué à établir.

F.2.2 Echographie longitudinale

Les tableaux F.3 présentent les résultats obtenus avec les échographies longitudinales. Les simulations ont été effectuées dans les mêmes conditions (même paramétrage et même prétraitement). Hormis pour le patient P18, nous pouvons constater que chaque patient est représenté par au moins trois images. Les patients P08, P10 et P11 ne possèdent pas d'acquisition (échographie et élastographie) en mode longitudinal d'où leur absence des tableaux. De prime abord, nous pouvons déjà noter que les échographies du patient P19 sont assez hétéroclites : les images n°223 et n°323 sont beaucoup plus sombres que les quatre autres. Ces quatre images ont été prises au niveau de la tête du thrombus alors que les deux autres ont été prises à quelques centimètres de la tête.

De plus, quel que soit la méthode de classification, les images d'un même patient ne sont pas toujours classées ensemble compliquant ainsi grandement l'interprétation des résultats. Globalement, nous retrouvons les grandes tendances constatées précédemment sur les échographies transversales : la CS directement sur les images trie celles-ci plutôt en fonction de leur échogénicité (les plus sombres d'un côté, les plus lumineuses d'un autre coté et les images intermédiaires) alors que la CS sur les coefficients du SO s'intéresse davantage à la texture des thrombi (peu ou pas de motifs échogènes, de gros motifs bien échogènes et contrastés, de nombreux petits motifs assez échogènes).

Classement selon l'origine de la thrombose veineuse profonde en mode longitudinal																													
Ori.	Cancer						C. EP	Immobilisation et EP						Génétique et EP						Idiopathique									
Pati.	P12						P18	P19						P20						P03					P01				
Age	56						95	51						91						37					88				
Gain	73						79	78						73						87					73				
Img.	004	008	011	015	018	021	028	031	033	035	022	023	027	123	127	223	323	032	132	232	332	007	009	015	107	011	013	015	016

(a) Classes idéales

Classification spectrale appliquée aux échographies à la taille 64x64 (dctpad) en mode longitudinal																													
2 Cl.	1											2		1											2				
3 Cl.	1											2											3						
4 Cl.	1											2											3		4				
Ori.	Immob. EP			Génétique et EP			Id.	Immob. EP			Cancer			C. EP	Cancer			Im. EP	Génétique et EP			Idiopathique		Im. EP	Ca.				
Pat.	P19			P20			P01	P19			P12	P18	P13			P21	P19	P03			P01		P19	P12					
Age	51			91			88	51			61	56	95			81	61	51	37			88		51	56				
Gain	78			73			73	78			93	73	79			88	93	78	87			73		78	73				
Img.	223	323	032	232	332	132	016	027	123	033	035	004	008	022	028	015	018	021	031	023	107	007	009	015	011	013	015	127	011

(b) Classification spectrale des images elles-mêmes

Classification spectrale appliquée aux scattering coefficients des échographies à la taille 64x64 (dctpad) en mode longitudinal																													
2 Cl.	1											2																	
3 Cl.	1											2																	
4 Cl.	1											2											3		4				
Ori.	Génét. EP		Immob. EP		Génét. EP		Ca.	Immob. EP		C. EP	Immob. EP		Génétique et EP			Idiopathique			Cancer										
Pat.	P20		P19		P20		P21	P19		P18	P19		P03			P01			P12						P13		P21		
Age	91		51		91		61	51		95	51		37			88			56						81		61		
Gain	73		78		73		93	78		79	78		87			73			73						88		93		
Img.	032	232	223	323	332	132	033	027	123	022	127	023	007	009	107	015	013	015	016	011	011	004	008	028	015	018	021	031	035

(c) Classification spectrale des coefficients du scattering operator

Tableau F.3 – Classification des échographies longitudinales en 2, 3 et 4 sous-ensembles.

F.3 Dureté du thrombus

F.3.1 Elastographie transversale

A présent, nous ne nous intéressons plus à l'échostructure du thrombus mais à sa dureté. Les résultats de classification en deux, trois et quatre sous-ensembles sont donnés au tableau F.4. Nous pouvons directement remarquer dans le tableau F.4.a que le thrombus du patient P19 est beaucoup plus dur (rouge) que les autres thrombus laissant penser que le thrombus est assez ancien. La classification spectrale seule (tableau F.4.b) semble classer les thrombi en fonction de leur dureté globale ce qui explique que le patient P19 soit isolé.

Le SO (avec le logarithme et avec le coefficient d'approximation) donne des sous-ensembles davantage différenciables par leur texture puisqu'il permet de dissocier les thrombi homogènes des thrombi hétérogènes. Nous pouvons également apercevoir que les élastographies des patients P13, P20 et P21 présentent des irrégularités (zones d'ombre que les ondes de cisaillement n'ont pas pu atteindre) et que, par conséquent, leurs images sont classées ensemble (quel que soit la méthode). Le lien entre facteurs déclencheurs et les classements des images paraît compliquer à établir. Par ailleurs, les classes obtenues sur les élastographies sont très différentes de celles sur les échographies laissant à penser que l'échogénicité du thrombus et sa dureté sont deux informations décorréées (sans lien entre elles).

Classement selon l'origine de la thrombose veineuse profonde en mode transversal															
Origine	Cancer				Ca. EP			Immob.	Im. EP	Ge. EP			Idiopa.		Id. EP
Patient	P12	P13	P21		P11	P18		P08	P19	P20	P03		P01	P10	
Age	56	81	61		93	95		79	51	91	37		88	57	
Gain	73	88	93		93	79		77	78	73	87		73	84	
	001	011	020	024	016	019	020	001	019	029	001	005	008	010	005
Image															

(a) Classes idéales

Classification spectrale appliquée aux élastographies à la taille 64x64 (dctpad) en mode transversal																				
2 Classes	1							2												
3 Classes	1					2					3									
4 Classes	1				2				3				4							
Origine	Génét. EP				Cancer				Id. EP	Génét. EP				Immob.	Cancer et EP			Génét. EP	Idiopa.	Im. EP
Patient	P20	P13	P21	P12	P10	P03	P01	P08	P18	P11	P03	P01	P19	P13	P21					
Age	91	81	61	56	57	37	88	79	95	93	37	88	81	61						
Gain	73	88	93	73	84	87	73	77	79	93	87	73	78	84						
	029	011	020	024	001	005	005	008	001	019	020	016	001	010	019					
Image																				

(b) Classification spectrale des images elles-mêmes

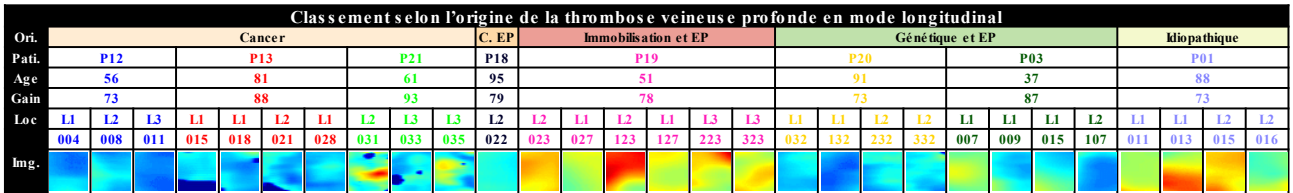
Classification spectrale appliquée aux scattering coefficients des élastographies à la taille 64x64 (dctpad) en mode transversal																		
2 Classes	1							2										
3 Classes	1					2					3							
4 Classes	1				2				3				4					
Origine	Cancer	Immob.	Id. EP	Idiopathique				Génét. EP	Cancer et EP				Génétiq. et EP			Im. EP	Cancer	
Patient	P12	P08	P10	P01				P03	P11	P18				P03	P20	P19	P13	P21
Age	56	79	57	88				37	93	95				37	91	51	81	61
Gain	73	77	84	73				87	93	79				87	73	78	88	93
	001	001	005	008	010	005	016	020	019	001	029	019	011	020	024			
Image																		

(c) Classification spectrale des coefficients du scattering operator

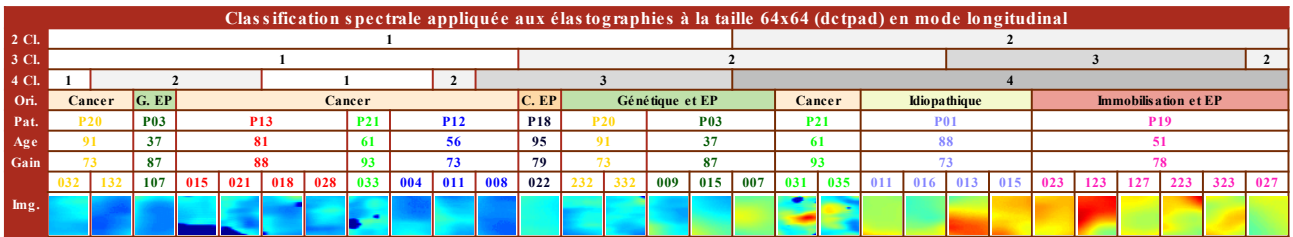
Tableau F.4 – Classification des élastographies transversales en 2, 3 et 4 sous-ensembles.

F.3.2 Elastographie longitudinale

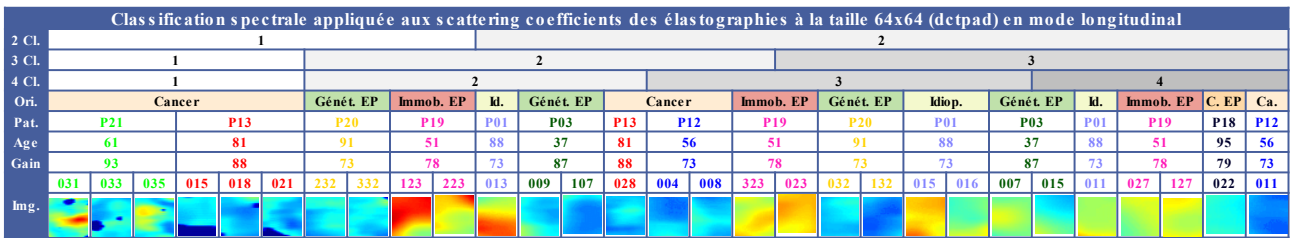
L'analyse des élastographies en mode longitudinal (voir figure F.5) donne des résultats globalement similaires : la classification spectrale trie les thrombi en fonction de leur dureté alors que le scattering operator caractérise davantage leur homogénéité. Les images du patient P19 sont beaucoup plus similaires en terme de dureté que d'échogénicité par rapport à la zone du thrombus analysée. Elles sont d'ailleurs dans le même sous-ensemble avec la classification spectrale. Les images d'un même patient restent, comme en échographie, pas toujours classées dans le même sous-ensemble. Enfin, l'épidémiologie des patients paraît difficile à retrouver à partir du classement des images.



(a) Classes idéales



(b) Classification spectrale des images elles-mêmes



(c) Classification spectrale des coefficients du scattering operator

Tableau F.5 – Classification des élastographies longitudinales en 2, 3 et 4 sous-ensembles.

F.4 Echogénicité et dureté du thrombus

Cette section combine les descripteurs des échographies et élastographies en les fusionnant : pour une image échographique donnée, son vecteur descripteur (image vectorisée ou coefficients du SO) est juxtaposé avec celui de l'image élastographique provenant de la même acquisition.

F.4.1 Echographie et élastographie transversales

Les simulations de cette section prennent en considération à la fois l'échogénicité et la dureté des thrombi. Le tableau F.6.b donne les résultats de la classification spectrale appliquée sur les vecteurs composés de la juxtaposition des échographies et des élastographies. Le patient P19 et son thrombus ancien est isolé comme en élastographie transversale. Ensuite, les sous-ensembles sont assez variables en fonction du nombre de classes (*e.g.* P18, P12 et P20 qui sont dans la classe 1 puis 2, puis 3). Nous pouvons également remarquer que les patients P01, P03 et P11 sont à nouveau associés (comme pour l'élastographie seule). Il en est de même des patients P10 (immobilisation), P13 (cancer), P21 (cancer), P12 (cancer) et P20 (idiopathique avec antécédents familiaux). Malgré la normalisation des deux types d'image, l'élastographie semble plus influencer sur les résultats de classification qui sont plus proches de ceux de l'élastographie seule que de l'échographie seule.

Le tableau F.6.c présente le classement des vecteurs descripteurs formés des coefficients du scattering operator des élastographies et des échographies (fusion en aval du scattering operator et en amont de la classification spectrale). Les résultats sont à nouveau proches de ceux obtenus sur l'élastographie seule puisque les patients P01, P10, P08, P12 et les patients P13, P19, P20, P21 sont regroupés. Que ce soit l'une ou l'autre des approches de classification, le classement des images ne paraît pas connecté aux facteurs déclencheurs ou à la présence d'EP.

Classement selon l'origine de la thrombose veineuse profonde en mode transversal															
Origine	Cancer				Ca. EP				Immob.	Im. EP	Gé. EP			Idiopa.	Id. EP
Patient	P12	P13	P21		P11	P18			P08	P19	P20	P03		P01	P10
Age	56	81	61		93	95			79	51	91	37		88	57
Gain	73	88	93		93	79			77	78	73	87		73	84
Image	001	011	020	024	016	019	020	001	019	029	001	005	008	010	005

(a) Classes idéales

Classification spectrale appliquée aux échographies/élastographies à la taille 64x64 (dctpad) en mode transversal																
2 Classes	1					2					1					2
3 Classes	1										2					3
4 Classes	1				2					3			4			3
Origine	Id. EP	Cancer			Gé. EP	Immob.	Idiopathique	Gé. EP	Ca. EP			Cancer	Gé. EP	Im. EP		
Patient	P10	P13	P21		P03	P08	P01	P03	P11	P18		P12	P20	P19		
Age	57	81	61		37	79	88	37	93	95		56	91	51		
Gain	84	88	93		87	77	73	87	93	79		73	73	78		
Image	005	011	020	024	005	001	008	010	001	016	019	020	001	029	019	

(b) Classification spectrale des images elles-mêmes

Classification spectrale appliquée aux scattering coefficients des échographie/élastographies à la taille 64x64 (dctpad) en mode transversal																				
2 Classes	1					2														
3 Classes	1										2					3				
4 Classes	1				2					3			4							
Origine	Ca. EP	Cancer			Généétique et EP		Im. EP	Cancer et EP		Gé. EP	Id. EP	Immob.	Idiopathique	Cancer						
Patient	P18	P21	P13		P03	P20	P19	P18	P11	P03	P10	P08	P01	P12						
Age	95	61	81		37	91	51	95	93	37	57	79	88	56						
Gain	79	93	88		87	73	78	79	93	87	84	77	73	73						
Image	020	020	024	011	001	029	019	019	016	005	005	001	008	010	001					

(c) Classification spectrale des coefficients du scattering operator

Tableau F.6 – Classification des échographies fusionnées avec les élastographies en mode transversal.

F.4.2 Échographie et élastographie longitudinales

Comme en transversal, les résultats en mode longitudinal affichés en figure F.7 sont plus similaires à ceux obtenus en élastographie qu'en échographie. Le nombre plus important d'images par patient en longitudinal qu'en transversal complique l'interprétation des résultats : certains patients (P19 et P20 notamment) n'ont pas l'ensemble de leurs images classé dans le même sous-ensemble. Toutefois, avec la classification spectrale seule, les patients P12, P13 et P03 semblent plutôt proches. Il en est de même des patients P01 et P19. Avec le scattering operator, nous obtenons un classement plus basé sur la texture où les patients P13/P21 et P01/P03/P12/P19 semblent former deux sous-groupes distincts. Le lien avec les facteurs déclencheurs et une possible EP ne paraît toujours pas évident sur ces simulations.

Classement selon l'origine de la thrombose veineuse profonde en mode longitudinal																																
Ori.	Cancer								C. EP	Immobilisation et EP								Génétique et EP						Idiopathique								
Pati.	P12								P18	P19								P20						P03			P01					
Age	56								95	51								91						37			88					
Gain	73								79	78								73						87			73					
Loc	L1	L2	L3	L1	L1	L2	L1	L2	L3	L3	L2	L2	L1	L2	L1	L3	L3	L1	L1	L2	L2	L1	L1	L1	L2	L1	L1	L2	L1	L1	L2	L2
Img.	004	008	011	015	018	021	028	031	033	035	022	023	027	123	127	223	323	032	132	232	332	007	009	015	107	011	013	015	016			

(a) Classes idéales

Classification spectrale appliquée aux échographies/élastographies à la taille 64x64 (dctpad) en mode longitudinal																																					
2 Cl.	1								1								2																				
3 Cl.	1								2								3																				
4 Cl.	1								2								3																				
Ori.	Cancer								Génétique et EP								Ca.	C. EP	Ca.	Immobilisation EP				Ca.	Idiopathique				Immobilisation EP				Gé EP				
PaL	P12								P03								P20				P21	P18	P21	P19				P21	P01				P19				P03
Age	56								37								91				61	95	61	51				61	88				51				37
Gain	73								87								73				93	79	93	78				93	73				78				87
Img.	004	008	011	015	018	021	028	009	015	107	032	132	232	332	033	022	035	027	223	323	031	016	011	013	015	023	123	127	007								

(b) Classification spectrale des images elles-mêmes

Classification spectrale appliquée aux scattering coefficients des échographies/élastographies à la taille 64x64 (dctpad) en mode longitudinal																																		
2 Cl.	1								2								3																	
3 Cl.	1								2								3																	
4 Cl.	1								2								3																	
Ori.	Cancer								Im. EP				Génétique et EP				C. EP	Immobilisation et EP				Idiopathique				Génétique et EP				Cancer				
PaL	P13								P21				P19				P18	P19				P01				P03				P12				P13
Age	81								61				51				95	51				88				37				56				81
Gain	88								93				78				79	78				73				87				73				88
Img.	015	018	021	031	033	035	223	323	032	232	332	132	022	027	023	127	123	013	011	015	016	007	009	015	107	008	011	004	028					

(c) Classification spectrale des coefficients du scattering operator

Tableau F.7 – Classification des échographies fusionnées avec les élastographies en mode longitudinal.

F.5 Bilan des simulations sur les images de thrombi

Globalement, les résultats de l'ensemble des simulations est assez cohérent avec le classement que nous aurions pu faire en observant uniquement les images (sans connaissance sur le contexte d'apparition des thrombi). La classification spectrale appliquée directement sur les images est plutôt basé sur l'intensité et le contraste général des images. Le scattering operator (avec notamment l'utilisation du logarithme) discrimine les images plutôt en fonction de leur texture (*i.e.* de leurs motifs). Cette dernière approche semble donc plus pertinente en échographie où l'échelle des niveaux de gris est altérée par l'égalisation d'histogramme et les changements de paramètres d'acquisition entre patients.

Le nombre plus important d'images en mode longitudinal complique l'interprétation des résultats (images d'un patient pas dans le même sous-ensemble). Toutefois, si nous fusionnons les résultats en mode transversal et en longitudinal, nous constatons certaines similitudes entre les sous-ensembles, indiquées dans le tableau F.8. En échographie, nous pouvons noter que les groupes sont différents en fonction de l'algorithme alors qu'il y a des points communs en élastographie. En effet, deux patients victimes d'un cancer sont associés ainsi que deux TVP idiopathiques (P01 et P03). Des simulations sur un plus grand panel de patients seraient nécessaires pour déterminer le lien réel entre la dureté du thrombus et la présence de thrombus.

Type	Classification spectrale	Scattering operator
Échographie	P01 (<i>Idiop.</i>) / P03 (<i>Gé. EP</i>) / P13 (<i>Cancer</i>)	P12 (<i>Cancer</i>) / P13 (<i>Cancer</i>)
	P12 (<i>Cancer</i>) / P20 (<i>Gé. EP</i>) / P21 (<i>Cancer</i>)	P03 (<i>Gé. EP</i>) / P21 (<i>Cancer</i>)
Élastographie	P01 (<i>Idiop.</i>) / P03 (<i>Gé. EP</i>)	P13 (<i>Cancer</i>) / P21 (<i>Cancer</i>)
	P12 (<i>Cancer</i>) / P13 (<i>Cancer</i>) / P20 (<i>Gé. EP</i>)	P01 (<i>Idiop.</i>) / P03 (<i>Gé. EP</i>)

Tableau F.8 – Similitudes entre la classification des images en mode transversal et longitudinal : chaque ligne des cellules du tableau indique deux ou trois patients qui ont été groupés ensemble lors des simulations transversales et longitudinales.

Comme le montrent les simulations précédentes et le tableau F.9, les sous-ensembles obtenus avec, respectivement, les échographies et les élastographies sont assez différents, notamment avec la classification spectrale. En transversale, nous pouvons voir que la classification spectrale associe à nouveau les patients P12, P20 et P21. Le scattering operator regroupe également le patient P20 avec un patient possédant un cancer (P13). En longitudinal, les sous-groupes sont beaucoup plus disparates en fonction du type d’images. En l’état actuel des travaux, il paraît difficile d’identifier un facteur déclencheur en fonction des résultats de classification mais ces simulations sont limitées par le nombre de patients et par le manque de certitudes sur la définition des classes dite « idéales ». Par ailleurs, bien que ces simulations semblent donner des résultats plutôt négatifs, il est trop tôt pour affirmer que ces approches sont inappropriées au contexte de caractérisation des thrombi veineux. Les futures simulations sur les images acquises avec le protocole 3 ainsi que sur les échographies et les élastographies de glandes salivaires apporteront sûrement davantage de réponses à ce sujet.

Mode	Classification spectrale	Scattering operator
Transversal	P12 (<i>Cancer</i>) / P20 (<i>Gé. EP</i>) / P21 (<i>Cancer</i>)	P13 (<i>Cancer</i>) / P20 (<i>Gé. EP</i>) P01 (<i>Idiop.</i>) / P08 (<i>Immob.</i>)
Longitudinal		P01 (<i>Idiop.</i>) / P12 (<i>Cancer</i>)

Tableau F.9 – Similitudes entre la classification des échographies et élastographies : chaque ligne des cellules du tableau indique deux ou trois patients qui ont été groupés ensemble lors des simulations réalisées sur les deux modes d’imageries.

Conclusion

La base de thrombi, collectée avec le protocole 2B étant limitée en taille, nous avons décidé d’appliquer la classification spectrale à cette base car les simulations sur la base de référence ont montré que la classification spectrale est plus performante que l’algorithme des k -moyennes, qu’elle soit appliquée sur les images directement ou sur les descripteurs du scattering operator. L’avantage du scattering operator pour classer les échographies est que cet algorithme cherche d’abord à caractériser la texture de l’image. Le SO semble donc peu impacté par les modifications des niveaux de gris induites par les changements des paramètres d’acquisition et l’égalisation d’histogramme. Toutefois, quelle que soit l’approche utilisée, il paraît difficile de corrélérer les résultats de classification et l’épidémiologie des patients. Certains patients présentant un cancer sont régulièrement associés mais ce ne sont pas toujours les mêmes et ils ne sont jamais tous dans le même sous-ensemble. Avec les élastographies, les sous-ensembles sont assez différents de ceux obtenus avec les échographies. Ce phénomène pourrait signifier que l’échogénicité et la dureté sont deux types d’information totalement indépendants. En combinant les deux types d’images pour créer nos vecteurs de descripteurs, les résultats de classification sont proches de ceux produits par les élastographies. Malgré la normalisation des images, l’élastographie semble donc prendre plus de poids dans le choix du classement des images.

Annexe G

Simulations complémentaires sur la base de thrombi

Glossaire de l'annexe B

CS Classification Spectrale (*spectral clustering* en anglais)

EP Embolie Pulmonaire

FGS *Fondamental Gray Scale* : échographie formée à partir de la fondamentale des ultrasons réfléchis

SMA Classification par Sélection de Modèles Affines

THI Tissue Harmonic Imaging : échographie des tissus humains formée à partir des harmoniques des ultrasons réfléchis

TVP Thrombose Veineuse Profonde

Introduction

Cette annexe présente les résultats de simulations complémentaires à celles dont les résultats sont détaillés en section 5.3 : la base utilisée (celle du protocole 3) et le protocole expérimental (paramètres des descripteurs, des classifieurs et des simulations) sont identiques. Toutefois, dans cette annexe, nous allons rechercher des classes différentes aux quatre propositions émises à la section 5.3 : identification des facteurs déclencheurs, identification du caractère provoqué de la Thrombose Veineuse Profonde (TVP) et/ou détection de la présence d'Embolie Pulmonaire (EP). Dans notre base de données, nous avons constaté que certaines TVP sont associées à un cancer. La section G.1 propose de classer les images de thrombi en deux catégories dans le but de dissocier les thromboses liées à un cancer des autres thromboses. Ensuite, la section G.2 s'intéresse à l'incidence du tabac sur l'apparition des TVP comme l'évoque la section 1.1.2.3. Enfin, la section G.3 envisage l'impact des paramètres d'acquisition sur la classification des échographies.

G.1 Détection d'un cancer

Dans cette section, nous proposons d'étudier les résultats de classification en deux classes en fonction de la présence d'un cancer lors de la survenue d'une TVP. Le tableau G.1 donne les pourcentages de bonne classification moyens (sur dix tests) de la Classification Spectrale (CS) pour chaque jeu de paramètres et de descripteurs détaillé dans le protocole expérimental. Contrairement à la section 5.3, les résultats de la classification supervisée et des k -moyennes ne sont pas affichés car ils ne sont pas forcément plus significatifs.

Simulations / Descripteurs :				Images	SO	SO _{log}	C ₂	C ₃	C ₄	C _{1,2}	C _{2,4}	C _{1,4}	SO + C ₂	SO + C _{1,2}	SO + C _{2,4}	SO + C _{1,4}	SO _{log} + C ₂	SO _{log} + C _{1,2}	SO _{log} + C _{2,4}	SO _{log} + C _{1,4}
T	Img	Descrip.	N°																	
Echographie			S01	38	51	64	50	61	69	49	55	56	49	51	49	49	64	64	54	54
			F S02	59	58	58	61	59	66	63	48	51	59	59	60	61	64	64	62	62
		N	S03	38	53	50	54	54	54	54	59	59	51	53	58	59	49	48	59	59
		N	F S04	59	55	54	51	64	53	51	55	55	47	43	56	54	55	54	54	55
		N	S05	51	56	49	56	61	69	54	53	56	51	51	51	51	46	47	58	59
		N	F S06	54	45	48	53	58	66	51	48	49	50	73	44	47	49	49	50	50
		N	N S07	51	52	50	54	54	54	54	59	59	51	53	58	58	49	50	59	59
		N	N F S08	54	57	51	51	61	55	51	56	56	46	48	54	55	57	53	55	56
Echographie EH			S09	38	51	64	50	61	68	49	55	56	49	51	49	49	64	64	54	54
			F S10	59	58	58	61	56	67	63	48	51	59	59	60	61	64	64	62	62
		N	S11	38	53	50	54	54	54	54	59	59	51	52	59	58	49	44	59	59
		N	F S12	59	56	52	51	56	57	51	57	55	47	45	55	55	55	54	55	54
		N	S13	51	56	49	55	61	69	55	53	55	51	51	51	47	45	59	59	
		N	F S14	53	44	50	52	58	66	51	49	48	50	73	46	47	49	49	50	50
		N	N S15	51	52	50	54	54	54	54	59	59	51	53	58	58	49	50	59	59
		N	N F S16	54	57	51	51	61	54	51	56	56	46	48	54	55	57	53	55	56
Elastographie			S17	60	59	60	57	61	74	59	62	64	59	62	63	63	59	59	59	59
			F S18	55	62	59	56	58	72	53	50	57	56	60	57	59	74	57	56	57
		N	S19	60	46	60	49	60	69	53	67	67	47	50	67	67	57	56	69	69
		N	F S20	56	55	59	42	57	69	43	68	71	46	50	68	69	58	58	69	71
		N	S21	60	61	54	55	62	74	58	53	49	53	54	47	47	56	56	58	57
		N	F S22	57	56	56	52	61	69	52	54	53	47	49	48	74	55	55	56	55
		N	N S23	60	46	59	54	69	55	69	69	46	49	69	69	58	56	69	69	
		N	N F S24	55	57	58	38	62	69	44	70	70	47	48	69	69	59	58	69	69

Tableau G.1 – Classification spectrale (avec l’algorithme des k -moyennes) des thrombi suivant leur association, ou non, à un cancer (uniformisation globale par ordre).

A travers le tableau G.1, nous observons qu’un certain nombre de simulations permettent d’atteindre les 70 % de bonne classification, aussi bien avec les échographies que les élastographies. Néanmoins, si nous nous penchons plus en détail sur les simulations donnant ces taux assez élevés (*cf.* les matrices de confusion du tableau G.2), nous constatons que la totalité des thrombi est regroupée en une seule classe. En s’attardant sur les meilleures simulations de la classification supervisée (*cf.* les matrices de confusion du tableau G.3), nous retrouvons des résultats similaires : la Sélection de Modèles Affines (SMA) ne parvient pas à modéliser correctement les classes et l’ensemble des images sont ensuite regroupées dans une même classe. Par conséquent, ces résultats suggèrent que la présence d’un cancer n’aurait pas d’impact direct sur la structure des thrombi, ou que cet impact ne serait pas visible par échographie et élastographie.

C _{cs}	Sans cancer	Avec cancer
Sans cancer	100.0	0.0
Avec cancer	100.0	0.0

(a) S21 - C₄ (74.4 %)

C _{es}	Sans cancer	Avec cancer
Sans cancer	100.0	0.0
Avec cancer	100.0	0.0

(b) S06 - SO + C_{1,2} (73.0 %)

Tableau G.2 – Matrice de confusion par classification spectrale en fonction de la présence d’un cancer.

C _{SMA}	Sans cancer	Avec cancer
Sans cancer	100.0	0.0
Avec cancer	98.0	2.0

(a) S24 - C_{2,4} (74.2 %)

C _{acp}	Sans cancer	Avec cancer
Sans cancer	100.0	0.0
Avec cancer	96.0	4.0

(b) S03 - SO + C_{1,2} (73.3 %)

Tableau G.3 – Matrice de confusion par SMA en fonction de la présence d’un cancer.

G.2 Incidence du tabagisme

La section 1.1.2.3 évoque des études [15, 26, 64] qui montreraient une incidence de la cigarette sur les thromboses veineuses. Cette section propose d'étudier les résultats de classification par rapport à la prise de tabac par les patients qui ont contracté une TVP. Les résultats par SMA, k -moyennes et Classification Spectrale (CS) (affichés au tableau G.4) ne dépassent pas les 66 % de bonne classification. L'étude des matrices de confusion de certaines simulations montre que, pour la majorité d'entre elles, les thrombi sont tous regroupés dans une même classe (tableaux G.5.a par CS, G.6.a et G.6.b par SMA).

Simulations / Descripteurs :				Images	SO	SO _{log}	C ₂	C ₃	C ₄	C _{1,2}	C _{2,4}	C _{1,4}	SO + C ₂	SO + C _{1,2}	SO + C _{2,4}	SO + C _{1,4}	SO _{log} + C ₂	SO _{log} + C _{1,2}	SO _{log} + C _{2,4}	SO _{log} + C _{1,4}	
T	Img	Descrip.	N°																		
Echographie			S01	49	54	50	56	55	56	57	55	56	55	56	57	57	51	50	57	57	
			S02	57	49	51	49	53	56	51	54	57	49	49	54	55	52	53	53	53	
		N	S03	49	53	58	49	46	57	49	43	43	46	47	45	44	52	54	43	43	
		N	F	S04	57	53	54	55	51	55	55	53	55	51	52	53	50	53	53	53	
		N		S05	62	45	55	49	55	55	52	53	56	49	49	51	51	56	55	50	51
		N	F	S06	65	59	61	61	54	56	59	54	54	58	62	58	61	59	61	64	64
		N	N	S07	62	52	58	49	46	57	49	43	43	46	48	45	44	52	55	43	43
		N	N	F	S08	65	51	52	52	51	54	54	52	52	54	53	49	53	49	53	52
Echographie EH			S09	49	54	50	56	55	56	57	55	56	55	56	57	57	51	50	57	57	
			S10	57	49	51	49	55	56	51	54	57	49	49	54	55	52	53	53	53	
		N	S11	49	53	58	49	46	57	49	43	43	46	47	44	45	52	52	43	43	
		N	F	S12	57	52	53	56	51	52	57	52	53	55	53	53	53	50	53	53	53
		N		S13	62	46	54	51	56	56	50	53	55	49	49	51	51	55	58	51	51
		N	F	S14	64	57	63	61	54	56	59	54	54	58	62	59	61	59	61	64	64
		N	N	S15	62	52	58	49	46	57	49	43	43	46	48	45	44	52	55	43	43
		N	N	F	S16	65	51	52	52	51	54	54	52	52	54	56	53	53	49	53	53
Elastographie			S17	55	54	47	51	48	64	54	55	57	49	51	52	53	51	48	48	49	
			S18	53	54	48	51	47	66	55	58	66	52	53	56	59	64	51	51	50	
		N	S19	55	54	60	56	56	52	58	51	51	52	55	51	51	56	56	54	54	
		N	F	S20	54	50	46	53	45	56	51	55	60	55	55	55	56	50	50	58	60
		N		S21	53	54	49	61	49	64	55	58	59	57	56	53	53	44	44	58	57
		N	F	S22	56	54	51	54	48	56	54	56	55	49	49	51	64	51	51	51	51
		N	N	S23	53	54	59	47	53	55	54	54	54	53	56	54	55	56	54	54	
		N	N	F	S24	57	52	47	52	49	56	53	59	59	55	55	56	56	49	50	58

Tableau G.4 – Classification spectrale (avec l'algorithme des k -moyennes) des thrombi par rapport au tabagisme (uniformisation globale par ordre).

Cependant, certaines simulations menées sur les échographies, comme celle du tableau G.5.b, partitionnent correctement les thrombi en fonction du tabagisme avec plus de 60 %. Cela pourrait indiquer que le tabac pourrait avoir une certaine incidence sur l'échogénicité du thrombus. Cependant, ces résultats sont assez précaires (peu de simulations affichant plus de 60 % de bonne classification) et il faudrait certainement une étude plus poussée sur l'impact du tabagisme avec des résultats positifs sur une plus grande quantité de données pour pouvoir étayer cette hypothèse.

C _{cs}	Non fumeur	Fumeur
Non fumeur	99.6	0.4
Fumeur	92.9	7.1

(a) S18 - C₄ (66.4 %)

C _{cs}	Non fumeur	Fumeur
Non fumeur	60.0	40.0
Fumeur	30.0	70.0

(b) S06 - SO_{log} + C_{2,4} (63.8 %)

Tableau G.5 – Matrice de confusion par classification spectrale par rapport au tabagisme.

C _{SMA}	Non fumeur	Fumeur
Non fumeur	100.0	0.0
Fumeur	97.1	2.9

(a) S18 - C₄ (64.2 %)

C _{SMA}	Non fumeur	Fumeur
Non fumeur	97.3	2.7
Fumeur	87.1	12.9

(b) S01 - C_{1,4} (64.4 %)

Tableau G.6 – Matrice de confusion par SMA par rapport au tabagisme.

G.3 Impact des paramètres d'acquisition des échographies

En étudiant plus en détail la base de données collectée avec le protocole 3 (*cf.* tableau 1.5), nous pouvons constater que l'ensemble des échographies n'a pas été obtenu avec les mêmes paramètres d'acquisition. En effet, deux modes d'échographie ont été utilisés : le mode fondamental (FGS) et le mode harmonique dit Tissue Harmonic Imaging (THI). La sous-section G.3.1 considère la classification des échographies en fonction de ces deux modes d'imagerie. Le tableau 1.5 montre également que le gain pouvait varier entre les acquisitions des différents patients pour permettre une meilleure visualisation du thrombus. La sous-section G.3.2 envisage l'influence du mode d'échographie et du gain sur la classification des thrombi.

G.3.1 Modes fondamental et harmonique

Le tableau 1.5 montre que les échographies ont été obtenues avec le mode FGS sur dix-sept patients et avec le mode THI sur dix-neuf patients.

Simulations / Descripteurs :				Images	SO	SO _{log}	c ₂	c ₃	c ₄	c _{1,2}	c _{2,4}	c _{1,4}	SO+ c ₂	SO+ c _{1,2}	SO+ c _{2,4}	SO+ c _{1,4}	SO _{log} + c ₂	SO _{log} + c _{1,2}	SO _{log} + c _{2,4}	SO _{log} + c _{1,4}
T	Img	Descrip.	N°																	
Echographie			S01	53	63	63	49	49	57	55	55	55	55	57	54	54	63	63	61	61
		F	S02	59	58	68	45	53	51	52	57	58	51	53	62	62	68	71	59	59
		N	S03	53	53	54	51	49	50	53	51	53	53	53	49	48	52	51	51	49
		N F	S04	59	49	52	54	54	49	53	49	48	50	50	52	52	54	55	47	47
		N	S05	50	53	62	49	49	57	53	54	54	53	52	55	55	57	57	57	57
		N	S06	56	46	58	47	53	50	47	56	56	50	49	53	53	56	58	52	52
		N N	S07	50	53	54	51	49	50	53	51	53	53	53	49	48	51	51	50	49
		N N F	S08	56	51	52	54	54	49	54	49	48	50	50	52	52	53	56	47	47
Echographie EH			S09	53	63	63	49	49	57	55	55	55	55	57	54	54	63	63	61	61
		F	S10	59	58	68	45	53	51	52	57	58	51	53	62	62	68	71	59	59
		N	S11	53	53	54	51	49	50	53	51	53	53	53	49	48	52	51	51	49
		N F	S12	59	49	52	54	54	49	53	49	48	50	50	52	52	54	55	47	47
		N	S13	50	53	62	49	49	57	53	54	54	53	52	55	55	57	57	57	57
		N	S14	56	46	58	47	53	50	47	56	56	50	49	53	53	56	58	52	52
		N N	S15	50	53	54	51	49	50	53	51	53	53	53	49	48	51	51	50	49
		N N F	S16	56	51	52	54	54	49	54	49	48	50	50	52	52	53	56	47	47

(a) SMA avec 50 % des thrombi consacrées à l'apprentissage (uniformisation globale par type)

Simulations / Descripteurs :				Images	SO	SO _{log}	c ₂	c ₃	c ₄	c _{1,2}	c _{2,4}	c _{1,4}	SO+ c ₂	SO+ c _{1,2}	SO+ c _{2,4}	SO+ c _{1,4}	SO _{log} + c ₂	SO _{log} + c _{1,2}	SO _{log} + c _{2,4}	SO _{log} + c _{1,4}
T	Img	Descrip.	N°																	
Echographie			S01	51	51	52	53	36	53	51	39	39	52	54	48	48	52	52	43	43
		F	S02	54	52	52	51	43	52	49	54	51	52	52	48	47	50	51	50	50
		N	S03	51	42	47	62	49	57	62	57	57	41	42	55	56	44	46	57	57
		N F	S04	54	61	51	45	55	55	45	53	53	59	61	52	53	53	56	53	53
		N	S05	57	39	51	42	36	53	44	36	39	43	43	46	46	51	51	49	48
		N	S06	56	45	49	50	43	52	49	48	49	55	46	58	55	46	46	50	50
		N N	S07	57	41	48	62	49	57	62	57	57	41	42	55	56	44	46	57	57
		N N F	S08	56	61	51	48	56	54	46	52	52	59	58	53	53	53	55	53	52
Echographie EH			S09	51	51	52	53	36	54	51	39	39	52	54	48	48	52	52	43	43
		F	S10	54	52	52	51	43	51	49	54	51	52	52	48	47	50	51	50	50
		N	S11	51	42	47	62	49	57	62	57	57	41	41	56	55	44	46	57	57
		N F	S12	54	61	51	44	58	52	43	52	53	59	60	53	53	53	56	53	53
		N	S13	57	40	51	42	37	53	43	36	39	43	43	46	46	51	51	48	48
		N	S14	55	44	49	50	43	52	49	49	48	56	46	57	55	46	46	50	50
		N N	S15	57	41	48	62	49	57	62	57	57	41	42	55	56	44	46	57	57
		N N F	S16	56	61	51	48	56	54	46	52	52	59	58	53	53	53	55	53	52

(b) Classification spectrale avec l'algorithme des k -moyennes (uniformisation globale par ordre)

Tableau G.7 – Classification des thrombi par rapport au mode d'imagerie échographique : FGS ou THI.

Les résultats de classification par SMA et par CS sont donnés par le tableau G.7. Ce dernier montre des taux globalement bas (aux alentours des 50 %) pour la majorité des simulations. Cependant, le tableau G.7.a montre que, en l'absence d'uniformisation des descripteurs par acquisition (S01, S02, S05, S06, S09, S10, S13 et S14), le mode d'échographie semble influencer la classification puisqu'environ 70 % des images sont bien classées en fonction des deux modes des coefficients du *scattering operator* (c'est moins le cas sur les multicorrélations). Cela se traduit également au niveau des matrices de confusion (*cf.* tableau G.8) où nous pouvons bien voir l'effet de l'uniformisation sur les descripteurs combinés : avec l'uniformisation, il n'est plus possible de classer les images en fonction des modes d'échographie (passage de 70.6 % à 50.6 %).

C_{SMA}	FGS	THI
FGS	73.8	26.3
THI	32.0	68.0

(a) $S02 - SO + C_{1,2}$ (70.6 %)

C_{SMA}	FGS	THI
FGS	42.5	57.5
THI	43.0	57.0

(b) $S03 - SO_{\log} + C_{1,2}$ (50.6 %)

Tableau G.8 – Matrice de confusion par SMA par rapport au mode d'imagerie échographique : fondamental ou harmonique.

Le tableau G.7.b semble indiquer que la classification spectrale est moins affectée par le changement de mode d'échographie puisqu'il y a au maximum 62 % de bonne classification par rapport aux deux modes (FGS et THI). Les simulations avec un taux proche des 50 %, comme celle dont le tableau G.9.b affiche la matrice de confusion suggère que le mode d'échographie n'a qu'une influence limitée sur le classement des images. Toutefois, lors de simulations plus défavorables, comme celles des tableaux G.9.a et G.9.c, les images du mode THI sont plutôt assez regroupées. Contrairement aux résultats par SMA, la classification spectrale donne de meilleurs résultats sans uniformisation des descripteurs par acquisition qu'avec cette uniformisation.

C_{cs}	FGS	THI
FGS	55.9	44.1
THI	35.0	65.0

(a) $S04 - SO$ (60.8 %)

C_{cs}	FGS	THI
FGS	50.0	50.0
THI	47.5	52.5

(b) $S04 - SO_{\log}$ (51.3 %)

C_{cs}	FGS	THI
FGS	52.9	47.1
THI	30.0	70.0

(c) $S03 - C_{1,2}$ (62.2 %)

Tableau G.9 – Matrice de confusion par classification spectrale par rapport au mode d'imagerie échographique : fondamental ou harmonique.

G.3.2 Modes d'imagerie et gain

Outre le mode d'échographie qui peut varier d'un patient à l'autre, il y a aussi le gain qui peut être ajusté pour améliorer la qualité de l'image. En fondamental (FGS), la majorité des acquisitions a été réalisée avec un gain de 75 dB (treize sur dix-sept). En harmonique, il y a un ajustement plus fréquent du gain : neuf patients avec un gain inférieur à 75 dB, neuf avec un gain de 75 dB et un seul avec un gain à 76 dB. Par conséquent, nous proposons d'étudier l'impact du mode et du gain en divisant la base en trois classes :

1. Mode FGS ;
2. Mode THI avec un gain inférieur à 75 dB ;
3. Mode THI avec un gain supérieur ou égal à 75 dB.

C_{SMA}	FGS	THI < 75	THI ≥ 75
FGS	85.0	13.8	1.3
THI < 75	60.0	36.7	3.3
THI ≥ 75	66.7	16.7	16.7

(a) $S05 - SO + C_{1,2}$ (55.9 %)

C_{SMA}	FGS	THI < 75	THI ≥ 75
FGS	85.0	13.8	1.3
THI < 75	60.0	36.7	3.3
THI ≥ 75	66.7	16.7	16.7

(b) $S13 - SO_{\log} + C_{1,2}$ (55.9 %)

Tableau G.10 – Matrice de confusion par SMA par rapport au mode d'imagerie échographique et au gain : harmonique avec un gain < 75 dB ou harmonique avec un gain ≥ 75 dB.

C_{cs}	FGS	THI < 75	THI \geq 75
FGS	51.8	30.6	17.6
THI < 75	13.1	57.7	29.2
THI \geq 75	0.0	35.7	64.3

(a) S08 - SO_{log} (56.2 %)

C_{cs}	FGS	THI < 75	THI \geq 75
FGS	52.9	33.5	13.5
THI < 75	14.6	85.4	0.0
THI \geq 75	54.3	24.3	21.4

(b) S14 - $C_{1,2}$ (58.4 %)

Tableau G.11 – Matrice de confusion par classification spectrale par rapport au mode d'imagerie échographique et au gain : harmonique avec un gain < 75 dB ou harmonique avec un gain \geq 75 dB.

En supervisé, le tableau G.12.a donne des taux de classification logiquement plus bas que ceux du tableau G.7.a, mais nous pouvons observer des tendances assez similaires : les taux les plus élevées (environ 50 %) ont été obtenus sur les descripteurs multirésolutions (*Scattering Operator* (SO)) sans uniformisation par acquisition. En observant les matrices de confusion des deux simulations donnant le pourcentage le plus élevé (*cf.* tableau G.10), nous constatons que la majorité des images sont regroupées dans la même classe.

Simulations / Descripteurs :			Images	SO	SO_{log}	c_2	c_3	c_4	$c_{1,2}$	$c_{2,4}$	$c_{1,4}$	$SO+$ c_2	$SO+$ $c_{1,2}$	$SO+$ $c_{2,4}$	$SO+$ $c_{1,4}$	$SO_{log}+$ c_2	$SO_{log}+$ $c_{1,2}$	$SO_{log}+$ $c_{2,4}$	$SO_{log}+$ $c_{1,4}$		
T	Img	Descrip.	N°																		
Echographie			S01	47	49	52	40	44	36	41	42	43	41	42	48	48	53	52	51	51	
		F	S02	51	50	51	48	47	48	48	48	48	48	48	50	50	51	52	47	48	
		N	S03	47	44	48	36	43	35	39	39	41	41	42	41	41	45	46	44	43	
		N	F	S04	51	50	49	48	46	49	49	48	48	49	48	48	49	49	47	47	
		N		S05	45	48	51	36	44	36	39	42	43	39	41	45	45	49	56	54	54
		N	F	S06	49	47	49	48	47	49	48	48	48	48	47	48	49	49	48	46	46
		N	N	S07	45	44	47	36	43	35	39	39	41	41	42	41	41	45	46	44	43
		N	N	F	S08	49	50	49	48	46	49	49	47	48	48	49	47	48	49	47	47
Echographie EH			S09	47	49	52	40	44	36	41	42	43	41	42	48	48	53	52	51	51	
		F	S10	51	50	51	48	47	48	48	48	48	48	48	50	50	51	52	47	48	
		N	S11	47	44	48	36	43	35	39	39	41	41	42	41	41	45	46	44	43	
		N	F	S12	51	50	49	48	46	49	49	48	48	48	49	48	48	49	47	47	
		N		S13	45	48	51	36	44	36	39	42	43	39	41	45	45	49	56	54	54
		N	F	S14	49	47	49	48	47	49	48	48	48	48	47	48	49	49	48	46	46
		N	N	S15	45	44	47	36	43	35	39	39	41	41	42	41	41	45	46	44	43
		N	N	F	S16	49	50	49	48	46	49	49	47	48	48	49	47	48	49	47	47

(a) SMA avec 50 % des thrombi consacrées à l'apprentissage (uniformisation globale par type)

Simulations / Descripteurs :			Images	SO	SO_{log}	c_2	c_3	c_4	$c_{1,2}$	$c_{2,4}$	$c_{1,4}$	$SO+$ c_2	$SO+$ $c_{1,2}$	$SO+$ $c_{2,4}$	$SO+$ $c_{1,4}$	$SO_{log}+$ c_2	$SO_{log}+$ $c_{1,2}$	$SO_{log}+$ $c_{2,4}$	$SO_{log}+$ $c_{1,4}$		
T	Img	Descrip.	N°																		
Echographie			S01	53	27	49	35	41	32	51	36	36	30	32	41	43	49	49	46	46	
		F	S02	48	49	50	33	36	38	50	46	42	30	50	36	41	44	48	52	50	
		N	S03	53	28	24	47	31	51	39	40	40	46	42	34	34	30	30	34	35	
		N	F	S04	46	43	55	36	38	44	29	40	41	32	33	35	35	41	36	34	36
		N		S05	54	48	42	50	40	32	51	48	48	43	44	46	47	44	44	43	43
		N	F	S06	39	46	44	55	36	41	57	52	54	45	49	51	38	43	46	46	
		N	N	S07	54	28	25	47	31	50	39	39	41	44	44	34	34	30	30	34	34
		N	N	F	S08	42	42	56	34	38	44	30	42	32	34	35	33	37	40	36	39
Echographie EH			S09	53	27	49	35	39	34	51	36	36	30	32	41	43	49	49	46	46	
		F	S10	48	49	50	33	35	39	50	46	42	30	50	36	41	44	48	52	50	
		N	S11	53	28	24	47	31	50	39	40	40	41	46	34	34	30	30	34	35	
		N	F	S12	50	40	55	32	39	44	30	42	41	34	31	33	34	37	35	32	37
		N		S13	56	48	42	50	38	34	51	48	49	43	45	48	46	43	44	43	43
		N	F	S14	41	45	44	53	34	39	58	54	54	46	51	51	49	35	43	47	46
		N	N	S15	54	28	25	47	31	50	39	39	41	44	44	34	34	30	30	34	34
		N	N	F	S16	42	42	56	34	38	45	30	42	32	34	35	33	37	40	36	39

(b) Classification spectrale avec l'algorithme des k -moyennes (uniformisation globale par ordre)

Tableau G.12 – Classification spectrale avec k -moyennes des thrombi par rapport au mode d'imagerie échographique : fondamental, harmonique avec un gain < 75 dB ou harmonique avec un gain \geq 75 dB (uniformisation globale par ordre).

En non supervisé, la classification spectrale montre des résultats plus variés avec, dans l'ensemble, des taux tout de même assez faibles (entre 30 et 50 %). Comme dans la sous-section précédente, nous remarquons que les taux les plus élevés correspondent aux simulations où les descripteurs ont été uniformisés par acquisition (soit le contraire de la SMA). Les matrices de confusion de deux des cas les plus défavorables (*cf.* tableau G.11) présentent des valeurs assez importantes sur leurs diagonales. La simulation S08 sur les coefficients du SO montre que plus de 50 % des images peuvent être classées suivant les trois classes considérées ici. Sur les multicorrélations, la simulation S14 permet de bien

classer 85.4 % des échographies harmoniques avec un gain inférieur à 75 dB : cela pourrait indiquer que le changement de gain peut perturber la classification spectrale des multicorrélations.

G.3.3 Conclusion sur l'influence des paramètres d'acquisition

Globalement, si nous regardons les résultats de classification pour l'ensemble des simulations présentées dans cette section, nous constatons que les taux de classification en fonction des deux modes d'échographie (et du gain) sont relativement bas pour une classification en deux (ou trois) classes. Néanmoins, pour certaines simulations, nous avons remarqué qu'il était possible de classer plus de la moitié des images par rapport aux deux (ou trois) classes considérées dans cette section. Nous avons également observé que l'uniformisation des descripteurs par acquisition semble limiter l'influence des changements de paramètres d'acquisition par SMA mais qu'elle produisait l'effet inverse avec la classification spectrale. En conclusion, nous pouvons nous demander si les taux assez importants de certaines simulations sont réellement le fait des changements de paramètres ou simplement le fruit du hasard (paramètres de simulations, choix des descripteurs, initialisation des k -moyennes). Nous avons par exemple obtenu de très bon taux de classification des échographies de glandes salivaires malgré des changements de paramètres entre les acquisitions. Pour analyser le réel impact des paramètres d'acquisition, il faudrait construire une base de données dans laquelle les images de chaque patient auraient été collectées pour un ensemble de jeux de paramètres prédéfinis.

Conclusion

En section 5.3, nous proposons de classer les thrombi suivants les facteurs déclencheurs, suivant le caractère provoqué ou idiopathique des thromboses et/ou la présence d'embolie pulmonaire. Cette annexe est donc un prolongement des simulations présentées dans cette section 5.3. Dans cette annexe, nous nous sommes intéressés à trois problématiques : la recherche d'un cancer associé à la thrombose, l'impact du tabagisme sur la structure du thrombus et l'influence de changements dans les paramètres d'acquisitions sur la classification des échographies. Les résultats des simulations ont montré que la présence d'un cancer associé à la thrombose ne semble pas pouvoir être détectée par échographie et élastographie avec les approches que nous avons considérées dans cette étude (peut-être qu'il existe des descripteurs et des classifieurs plus pertinents qui le permettraient). Néanmoins, ces résultats ne sont pas forcément illogiques puisque, suivant la nature du cancer, ce sont pas forcément les mêmes mécanismes physiologiques qui sont à l'origine de la thrombose. En ce qui concerne le tabagisme, quelques simulations parviennent à classer correctement plus de la moitié des échographies en fonction de la prise de tabac (environ 60 % de bon classement). Toutefois, la majorité des simulations ont donné des résultats négatifs puisque la grande partie des thrombi ont été regroupés dans une même classe. Il faudrait donc une étude plus approfondie avec davantage de données pour être certains que le tabagisme ait un réel impact sur la structure du thrombus. Enfin, la classification des échographies en fonction des paramètres d'acquisition (mode d'échographie avec ou sans prise en compte du gain) ont donné, globalement, des pourcentages relativement faibles qui pourraient donc laisser à penser que leur impact est négligeable. Cependant, quelques simulations affichent un taux de classification assez élevé (70 % suivant les deux modes fondamental et harmonique, 55 % en incluant le gain), ce qui pourrait indiquer que la modification des paramètres d'acquisition peut induire un biais dans la classification des échographies. Les bons résultats sur les glandes salivaires montrent néanmoins que ce biais peut-être négligeable si les frontières entre les classes sont assez bien définies (*i.e.* si les classes sont suffisamment éloignées les unes des autres dans l'espace des descripteurs).

Bibliographie

- [1] A. Achim, A. Bezerianos, and P. Tsakalides. Novel bayesian multiscale method for speckle removal in medical ultrasound images. *IEEE Transactions on Medical Imaging*, 20(8) : 772–783, August 2001. Cité page 35
- [2] A. Alunno, E. Bartoloni, and R. Gerli. *Sjogren's Syndrome : Novel Insights in Pathogenic, Clinical and Therapeutic Aspects*. Elsevier Science, June 2016. Cité page 26
- [3] M. Ankerst, M.M. Breunig, H.P. Kriegel, and J. Sander. OPTICS : Ordering Points To Identify the Clustering Structure. In *ACM SIGMOD International Conference on Management of Data*, pages 49–60, Philadelphia, Pennsylvania, USA, June 1999. Cité page 38
- [4] A.F. Badea, M. Lupsor Platon, M. Crisan, C. Cattani, I. Badea, G. Pierro, G. Sannino, and G. Baciut. Fractal analysis of elastographic images for automatic detection of diffuse diseases of salivary glands : Preliminary results. *Computational and Mathematical Methods in Medicine*, 2013, April 2013. Cité page 27
- [5] C.N. Bagot and R. Arya. Virchow and his triad : a question of attribution. *British Journal of Haematology*, 143(2) : 180–190, September 2008. Cité page 4
- [6] A. Baille, P. Capéraà, and B. van Cutsem. *Méthodes et modèles en statistique non paramétrique, exercices et compléments*. Presses de l'Université Laval, January 1990. Cité page 182
- [7] A.K. Barros, A. Mansour, and N. Ohnishi. Removing artifacts from electrocardiographic signals using independent components analysis. *Neurocomputing*, 22(1–3) : 173–186, November 1998. Cité page 38
- [8] K. Bensafia, A. Mansour, and S. Haddab. Blind source subspace separation and classification of ecg signals. In *Conférence Internationale en Automatique & Traitement de Signal*, Sousse, Tunisia, March 2017. Cité page 80
- [9] T. Berthomier, A. Mansour, L. Bressollette, F. Le Roy, and D. Mottier. Deep venous thrombosis : Database creation and image preprocessing. In *International Conference on Frontiers of Signal Processing*, pages 87–92, Warsaw, Poland, October 2016. 2 citations pages 124 et 147
- [10] T. Berthomier, A. Mansour, L. Bressollette, F. Le Roy, and D. Mottier. Venous blood clot structure characterization using scattering operator. In *International Conference on Frontiers of Signal Processing*, pages 73–80, Warsaw, Poland, October 2016. 2 citations pages 124 et 146
- [11] T. Berthomier, A. Mansour, L. Bressollette, F. Le Roy, and D. Mottier. Deep venous thrombus characterization : ultrasonography , elastography and scattering operator. *Advances in Science, Technology and Engineering Systems Journal*, 2(3) : 48–59, April 2017. 2 citations pages 124 et 145
- [12] T. Berthomier, A. Mansour, L. Bressollette, F. Le Roy, D. Mottier, L. Frechier, and B. Hermenault. Scattering operator and spectral clustering for ultrasound images : Application on deep venous thrombi. In *International Conference on Biomedical Signal and Image Analysis*, pages 630–637, Paris, France, October 2017. World Academy of Science, Engineering and Technology. 4 citations pages 124, 146, 197 et 199
- [13] T. Berthomier, A. Mansour, L. Bressollette, D. Mottier, F. Le Roy, B. Hermenault, C. Hoffmann, and L. Frechier. Unsupervised clustering using high order statistics analysis for ultrasound images of thrombi [submitted : acceptance on october 13, 2018]. In *International Conference on Bioinformatics and Biomedicine*, Madrid, Spain, December 2018. 2 citations pages 124 et 145

- [14] S. Beuil. *Mémoire de DESC de médecine vasculaire : Recherche d'une corrélation entre l'échographie 2D et l'Elastométrie sur les thromboses veineuses proximales profondes pour datation du thrombus*. PhD thesis, UBO, Brest, France, June 2013. 2 citations pages 6 et 30
- [15] M. Blondon, K.L. Wiggins, B. McKnight, B.M. Psaty, K.M. Rice, S.R. Heckbert, and N.L. Smith. The association of smoking with venous thrombosis in women : a population-based, case-control study. *Thrombosis and Haemostasis*, 109(5) : 891–896, May 2013. 2 citations pages 6 et 209
- [16] J.L. Bosson, J. Labarere, M.A. Sevestre, J. Belmin, L. Beyssier, A. Elias, A. Franco, and P. Le Roux. Deep vein thrombosis in elderly patients hospitalized in subacute care facilities : a multicenter cross-sectional study of risk factors, prophylaxis, and prevalence. *Archives of Internal Medicine*, 163(21) : 2613–2618, November 2003. Cité page 5
- [17] J. Bouffartigue and A.M. Delrieu. *Les racines grecques*. Belin, December 1996. Cité page 1
- [18] E. Bournay Bouchereau. *Analyse d'images par transformées en ondelettes. Application aux images sismiques*. PhD thesis, Université Joseph-Fourier, Grenoble, France, March 1997. 3 citations pages 170, 173 et 174
- [19] L. Bressollette, P. Quehe, S. Gestin, and S. Beuil. Imagerie du thrombus. *Journal des Maladies Vasculaires*, 40(5) : 286, September 2015. Cité page 151
- [20] C. Brooker, I. Langlois-Wils, and E. Lepresle. *Le corps humain : étude, structure et fonction*. Infirmier(e)s : étudiants (Hors collection). De Boeck Supérieur, November 2000. Cité page 26
- [21] J. Bruna. *Scattering representations for recognition*. PhD thesis, Ecole Polytechnique X, Paris, France, November 2013. 2 citations pages 72 et 161
- [22] J. Bruna and S. Mallat. Invariant Scattering Convolution Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(8) : 1872–1886, August 2013. 12 citations pages 38, 45, 55, 72, 73, 74, 75, 76, 79, 87, 110 et 167
- [23] R. Cadene, T. Robert, N. Thome, and M. Cord. M2CAI workflow challenge : convolutional neural networks with time smoothing and hidden Markov model for video frames classification. *Computing Research Repository*, arXiv :1610.05541, December 2016. Cité page 49
- [24] J.F. Cardoso. Eigen-structure of the fourth-order cumulant tensor with application to the blind source separation problem. In *International Conference on Acoustics, Speech, and Signal Processing*, pages 2655–2658, Albuquerque, New Mexico , USA, April 1990. Cité page 84
- [25] J.F. Cardoso and Antoine Souloumiac. Blind beamforming for non-gaussian signals. *IEE Proceedings F - Radar and Signal Processing*, 140(6) : 362–370, December 1993. Cité page 84
- [26] P. Carruzzo, M. Méan, A. Limacher, D. Aujesky, J. Cornuz, and C. Clair. Association between smoking and recurrence of venous thromboembolism and bleeding in elderly patients with past acute venous thromboembolism. *Thrombosis Research*, 138 : 74–79, February 2016. 2 citations pages 6 et 209
- [27] L. Castéra, J. Vergniol, J. Foucher, B. Le Bail, E. Chanteloup, M. Haaser, M. Darriet, P. Couzigou, and V. de Lédinghen. Prospective comparison of transient elastography, Fibrotest, APRI, and liver biopsy for the assessment of fibrosis in chronic hepatitis C. *Gastroenterology*, 128(2) : 343–350, February 2005. Cité page 35
- [28] S.G. Chang, B. Yu, and M. Vetterli. Spatially adaptive wavelet thresholding with context modeling for image denoising. *IEEE Transactions on Image Processing*, 9(9) : 1522–1531, September 2000. Cité page 35
- [29] C. Chaux. *Analyse en ondelettes M-bandes en arbre dual; application à la restauration d'images*. PhD thesis, Université de Marne-la-Vallée, Marne-la-Vallée, France, December 2006. 2 citations pages 168 et 169
- [30] C.M. Chen, Y.H. Chou, K.C. Han, G.S. Hung, C.M. Tiu, H.J. Chiou, and S.Y. Chiou. Breast lesions on sonograms : computer-aided diagnosis with nearly setting-independent features and artificial neural networks. *Radiology*, 226(2) : 504–514, February 2003. 2 citations pages 37 et 38

- [31] H.D. Cheng, J. Shan, W. Ju, Y. Guo, and L. Zhang. Automated breast cancer detection and classification using ultrasound images : A survey. *Pattern Recognition*, 43(1) : 299–317, January 2010. *3 citations pages 35, 36 et 37*
- [32] S. Chevrel, A. Bellion, and S. Liem. La phlébite : 2 français sur 3 ne connaissent pas les 3 principaux symptômes. *Communiqué de presse de Bayer HealthCare*, March 2014. <https://www.bayer.fr/sites/default/files/2014-03-21-la-phlebite-3-signes-pour-reagir.pdf> [Online : accessed on December 22, 2015]. *2 citations pages 1 et 6*
- [33] S. Chiou, F. Forsberg, T.B. Fox, and L. Needleman. Comparing differential tissue harmonic imaging with tissue harmonic and fundamental gray scale imaging of the liver. *Journal of Ultrasound in Medicine*, 26(11) : 1557–1563, November 2007. *2 citations pages 12 et 13*
- [34] P. Chuzel, A. Mansour, J. Ognard, J. Gentric, L. Bressollette, D. Hamad, and N. Betrouni. Automatic clustering for MRI images, application on perfusion MRI of brain. In *International Conference on Frontiers of Signal Processing*, pages 63–66, Warsaw, Poland, October 2016. *3 citations pages 36, 50 et 51*
- [35] D. Ciresan, A. Giusti, L.M. Gambardella, and J. Schmidhuber. Deep Neural Networks Segment Neuronal Membranes in Electron Microscopy Images. In F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, editors, *Neural Information Processing Systems*, pages 2843–2851, Lake Tahoe, Nevada, USA, December 2012. *Cité page 38*
- [36] A.T. Cohen, G. Agnelli, F.A. Anderson, J.I. Arcelus, D. Bergqvist, J.G. Brecht, I.A. Greer, J.A. Heit, J.L. Hutchinson, A.K. Kakkar, D. Mottier, E. Oger, M.M. Samama, and M. Spannagl. Venous thromboembolism (VTE) in Europe. *Thrombosis and Haemostasis*, 98(4) : 756–764, September 2007. *Cité page 1*
- [37] P. Comon. Tensor diagonalization, a useful tool in signal processing. *IFAC Proceedings Volumes*, 27(8) : 77–82, July 1994. *Cité page 84*
- [38] P. Comon, P. Chevalier, and V. Capdevielle. Performance of contrast-based blind source separation. In *IEEE Signal Processing Advances in Wireless Communications*, pages 345–348, Paris, France, April 1997. *Cité page 84*
- [39] P. Comon and C. Jutten. *Handbook of Blind Source Separation : Independent Component Analysis and Applications*. Elsevier Science, February 2010. *3 citations pages 38, 80 et 84*
- [40] J. Constans, Société Française de Médecine Vasculaire, Collège des Enseignants de Médecine Vasculaire, and Collège Français de Pathologie Vasculaire. *Les explorations vasculaires*. Elsevier Health Sciences, January 2014. *5 citations pages 9, 11, 18, 19 et 151*
- [41] D. Cornec, V. Devauchelle-Pensec, A. Saraux, and S. Jousse-Joulin. Clinical usefulness of salivary gland ultrasonography in Sjögren’s syndrome : Where are we now? *La Revue de Médecine Interne*, 37(3) : 186–194, November 2016. *2 citations pages 27 et 108*
- [42] J.M. Correas, A.M. Tissier, A. Khairoune, G. Khoury, D. Eiss, and O. Hélénon. Elastographie ultrasonore de la prostate : état de l’art. *Journal de Radiologie Diagnostique et Interventionnelle*, 94(5) : 567–576, May 2013. *Cité page 35*
- [43] M. Crosier and L.D. Griffin. Texture classification with a dictionary of basic image features. In *Conference on Computer Vision and Pattern Recognition*, pages 1–7, Anchorage, Alaska, USA, June 2008. *2 citations pages 91 et 95*
- [44] A. Dahabiah, J. Puentes, B. Guias, L. Bressollette, and B. Solaiman. Comparative Neural Network Based Venous Thrombosis Echogenicity and Echostructure Characterization using Ultrasound Images. In *Information and Communication Technologies*, volume 1, pages 992–997, Damascus, Syria, April 2006. *Cité page 31*
- [45] A. Dahabiah, J. Puentes, and B. Solaiman. Venous Thrombosis Supervised Image Indexing and Fuzzy Retrieval. In *International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 4528–4531, Lyon, France, August 2007. *Cité page 31*
- [46] K.J. Dana, B. Van Ginneken, S.K. Nayar, and J.J. Koenderink. Reflectance and texture of real-world surfaces. *ACM Transactions On Graphics*, 18(1) : 1–34, January 1999. *3 citations pages 91, 93 et 94*

- [47] B. Daniel, D. Françoise, and A. Isabelle. *Rhumatologie*. Le livre de l'interne. Lavoisier, December 2014. *Cité page 26*
- [48] L. De Lathauwer, B. De Moor, and J. Vandewalle. Independent component analysis and (simultaneous) third-order tensor diagonalization. *IEEE Transactions on Signal Processing*, 49(10) : 2262–2271, October 2001. *Cité page 84*
- [49] Société Française de Médecine Vasculaire, Collège des enseignants de médecine Vasculaire, and Collège Français de Pathologie Vasculaire. *Traité de médecine vasculaire. Tome 2 : Maladies veineuses, lymphatiques et microcirculatoires, thérapeutique*. Elsevier Health Sciences, February 2012. *5 citations pages 6, 7, 8, 13 et 151*
- [50] D. Declercq and A. Quinquis. *Signaux et systèmes en questions : le signal aléatoire*. Hermes Science, September 1996. *5 citations pages 179, 180, 187, 189 et 190*
- [51] J. Di Carlo and D. Cox. Untangling invariant object recognition. *Trends in cognitive sciences*, 11(8) : 333–341, August 2007. *Cité page 37*
- [52] M. Ester, H.P. Kriegel, J. Sander, and X. Xu. Density-based spatial clustering of applications with noise. In *International Conference on Knowledge Discovery and Data Mining*, pages 226–231, Portland, Oregon, USA, August 1996. *2 citations pages 38 et 50*
- [53] H. Fanet. *Medical Imaging Based on Magnetic Fields and Ultrasounds*. Wiley, February 2014. *4 citations pages 9, 10, 14 et 151*
- [54] T. Ferré. *Traitements d'images optroniques*. Edition Techniques Ingénieur, June 1996. *Cité page 159*
- [55] D.H. Fisher. Knowledge acquisition via incremental conceptual clustering. *Machine learning*, 2(2) : 139–172, July 1987. *Cité page 38*
- [56] M. Fritz, E. Hayman, B. Caputo, and J. Eklundh. The kth-tips database. www.nada.kth.se/cvap/databases/kth-tips/, May 2004. [Online : accessed on August 23, 2018]. *5 citations pages xvi, 89, 91, 146 et 197*
- [57] J.P. Galanaud, A.C. Arnoult, M.A. Sevestre, C. Genty, M. Bonaldi, A. Guyard, P. Giordana, O. Pichot, M. Colonna, I. Quéré, and J.L. Bosson. Impact of anatomical location of lower limb venous thrombus on the risk of subsequent cancer. *Thrombosis and Haemostasis*, 112(06) : 1129–1136, December 2014. *2 citations pages 6 et 124*
- [58] B.S. Garra. Elastography : history, principles, and technique comparison. *Abdominal Imaging*, 40(4) : 680–697, April 2015. *Cité page 31*
- [59] B.S. Garra, B.H. Krasner, S.C. Horii, S. Ascher, S.K. Mun, and R.K. Zeman. Improving the distinction between benign and malignant breast lesions : the value of sonographic texture analysis. *Ultrasonic imaging*, 15(4) : 267–285, October 1993. *Cité page 37*
- [60] J.L. Gennisson, T. Deffieux, M. Fink, and M. Tanter. Ultrasound elastography : Principles and techniques. *Diagnostic and Interventional Imaging*, 94(5) : 487–495, May 2013. *2 citations pages 13 et 14*
- [61] A. Géron. *Machine Learning avec Scikit-Learn : Mise en oeuvre et cas concrets*. Dunod, August 2017. *3 citations pages 43, 45 et 46*
- [62] G.B. Giannakis and M.K. Tsatsanis. Signal detection and classification using matched filtering and higher order statistics. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 38(7) : 1284–1296, July 1990. *Cité page 80*
- [63] H. Goh, N. Thome, M. Cord, and J. Lim. Learning deep hierarchical visual feature coding. *IEEE Transactions on Neural Networks and Learning Systems*, 25(12) : 2212–2225, March 2014. *Cité page 48*
- [64] B.A. Golomb, V.T. Chan, J.O. Denenberg, S. Koperski, and M.H. Criqui. Risk marker associations with venous thrombotic events : a cross-sectional analysis. *British Medical Journal*, 4(3) : e003208, March 2014. *2 citations pages 6 et 209*
- [65] I. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning*. MIT Press, November 2016. *2 citations pages 46 et 48*

- [66] S. Guha, R. Rastogi, and K. Shim. Rock : A robust clustering algorithm for categorical attributes. In *International Conference on Data Engineering*, pages 512–521, Sydney, New South Wales, Australia, March 1999. *Cité page 38*
- [67] Y. Guo, H.D. Cheng, J. Huang, J. Tian, W. Zhao, L. Sun, and Y. Su. Breast ultrasound image enhancement using fuzzy logic. *Ultrasound in medicine & biology*, 32(2) : 237–247, February 2006. *Cité page 36*
- [68] Iijima H. *Medical Review : Approaches to the Diagnosis of Liver Fibrosis*. Toshiba Medical Systems Corporation, September 2014. <https://medical.toshiba.com/download/ul-aplio-500-TMR-Approaches-to-the-Diagnosis-of-Liver-Fibrosis.pdf> [Online : accessed on October 16, 2017]. *2 citations pages 16 et 17*
- [69] M.E. Hachemi, S. Callé, and J.P. Remeniéras. Utilisation des ondes de cisaillement ultrasonores pour l’imagerie d’élasticité des tissus biologiques. *Traitement du signal et cancérologie*, 23(3-4-NS) : 247–258, December 2006. *2 citations pages 9 et 13*
- [70] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I.H. Witten. The WEKA data mining software : an update. *ACM SIGKDD explorations newsletter*, 11(1) : 10–18, June 2009. *Cité page 38*
- [71] D. Hamad and P. Biela. Introduction to spectral clustering. In *International Conference on Information and Communication Technologies : From Theory to Applications*, pages 1–6, Damascus, Syria, April 2008. *2 citations pages 39 et 51*
- [72] J.A. Hartigan and M.A. Wong. Algorithm AS 136 : A k-means clustering algorithm. *Journal of the Royal Statistical Society*, 28(1) : 100–108, January 1979. *3 citations pages 38, 49 et 55*
- [73] J. Herault and C. Jutten. *Réseaux neuronaux et traitement du signal*. Hermes Science, July 1994. *Cité page 47*
- [74] K. Horsch, M.L. Giger, L.A. Venta, and C.J. Vyborny. Automatic segmentation of breast lesions on ultrasound. *Medical Physics*, 28(8) : 1652–1659, August 2001. *Cité page 36*
- [75] Y.L. Huang, K.L. Wang, and D.R. Chen. Diagnosis of breast tumors with ultrasonic texture analysis using support vector machines. *Neural Computing & Applications*, 15(2) : 164–169, April 2006. *2 citations pages 37 et 38*
- [76] INSERM. Dossier d’information : Imagerie médicale. <http://www.inserm.fr/thematiques/technologies-pour-la-sante/dossiers-d-information/imagerie-fonctionnelle-biomedicale>, October 2011. [Online : accessed on January 23, 2017]. *Cité page 34*
- [77] F. Jobin. *La thrombose*. Presses de l’Université Laval, February 1995. *5 citations pages 4, 5, 6, 8 et 149*
- [78] I.T. Jolliffe. *Principal component analysis*. Springer, April 2002. *Cité page 41*
- [79] S. Joo, Y.S. Yang, W.K. Moon, and H.C. Kim. Computer-aided diagnosis of solid breast nodules : use of an artificial neural network based on multiple sonographic features. *IEEE Transactions on Medical Imaging*, 23(10) : 1292–1300, October 2004. *2 citations pages 36 et 37*
- [80] S. Jousse-Joulin, V. Milic, M.V. Jonsson, A. Plagou, E. Theander, N. Luciano, P. Rachele, C. Baldini, H. Bootsma, A. Vissink, A. Hocevar, S. De Vita, Z. Tzioufas, A.G. and Alavi, S.J. Bowman, and V. Devauchelle-Pensec. Is salivary gland ultrasonography a useful tool in sjögren’s syndrome? a systematic review. *Rheumatology*, 55(5) : 789–800, May 2016. *3 citations pages 27, 35 et 108*
- [81] Y.M. Kadah, A.A. Farag, J.M. Zurada, A.M. Badawi, and A.B.M. Youssef. Classification algorithms for quantitative tissue characterization of diffuse liver disease from ultrasound images. *IEEE Transactions on Medical Imaging*, 15(4) : 466–478, August 1996. *3 citations pages 37, 38 et 55*
- [82] M. Kass, A. Witkin, and D. Terzopoulos. Snakes : Active contour models. *International Journal of Computer Vision*, 1(4) : 321–331, January 1988. *Cité page 36*

- [83] A. Knopf, B. Hofauer, K. Thürmel, R. Meier, K. Stock, M. Bas, and N. Manour. Diagnostic utility of Acoustic Radiation Force Impulse (ARFI) imaging in primary sjogren's syndrome. *European Radiology*, 25(10) : 3027–3034, October 2015. *3 citations pages 108, 121 et 137*
- [84] T. Kohonen. The self-organizing map. *Proceedings of the IEEE*, 78(9) : 1464–1480, September 1990. *Cité page 38*
- [85] C. Kollmann. New sonographic techniques for harmonic imaging—Underlying physical principles. *European Journal of Radiology*, 64(2) : 164–172, November 2007. *Cité page 12*
- [86] A. Krizhevsky, I. Sutskever, and G.E. Hinton. Imagenet classification with deep convolutional neural networks. In *Neural Information Processing Systems*, pages 1097–1105, Lake Tahoe, Nevada, USA, December 2012. *2 citations pages 47 et 48*
- [87] J.B. Kruskal. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, 29(1) : 1–27, March 1964. *2 citations pages 43 et 55*
- [88] J.L. Lacoume, P.O. Amblard, and P. Comon. *Statistiques d'ordre supérieur pour le traitement du signal*. Elsevier Science, October 1997. *5 citations pages 80, 186, 187, 192 et 193*
- [89] R.S. Lazebnik. *Whitepaper : Tissue Strain Analytics Virtual Touch Tissue Imaging and Quantification*. Siemens Medical Solutions, October 2008. <https://usa.healthcare.siemens.com/ultrasound/tissue-strain-analytics/virtual-touch-tissue-quantification> [Online : accessed on August 18, 2017]. *Cité page 15*
- [90] Y. LeCun, L. D. Jackel, L. Bottou, C. Cortes, J. S. Denker, H. Drucker, I. Guyon, U. A. Muller, E. Sackinger, P. Simard, and V. Vapnik. Learning algorithms for classification : A comparison on handwritten digit recognition. In *Neural Networks : The Statistical Mechanics Perspective*, pages 261–276, Pohang, Korea, February 1995. *Cité page 80*
- [91] T.H. Lee, M.F.A. Fauzi, and R. Komiya. Segmentation of ct brain images using k-means and em clustering. In *International Conference on Computer Graphics, Imaging and Visualisation*, pages 339–344. Penang, Malaysia, August 2008. *2 citations pages 36 et 38*
- [92] W.L. Lee, T. Tan, T. Falkmer, and Y. H. Leung. Single-trial event-related potential extraction through one-unit ica-with-reference. *Journal of neural engineering*, 13(6) : 066010, October 2016. *Cité page 38*
- [93] P. Lemberger, M. Batty, M. Morel, and J.L. Raffaëlli. *Big Data et Machine Learning - Les concepts et les outils de la data science*. Dunod, October 2016. *4 citations pages 45, 46, 47 et 55*
- [94] V.P. Leonov and A.N. Shiryaev. On a method of calculation of semi-invariants. *Theory of Probability & Its Applications*, 4(3) : 319–329, March 1959. *Cité page 180*
- [95] P. Léger, D. Barcat, C. Boccalon, J. Guilloux, and H. Boccalon. Thromboses veineuses des membres inférieurs et de la veine cave inférieure. *EMC - Cardiologie-Angéiologie*, 1(1) : 80–96, February 2004. *5 citations pages 4, 5, 6, 150 et 151*
- [96] P.H. Lim, R. Debon, B. Solaiman, L. Bressollette, B. Guias, C. Roux, and H.T. Chuah. Toward determination of venous thrombosis ages by using fuzzy logic and supervised Bayes classification. In *International Conference of the IEEE Engineering in Medicine and Biology Society*, volume 4, pages 3831–3834, Istanbul, Turkey, October 2001. *Cité page 31*
- [97] Paweł Liskowski and Krzysztof Krawiec. Segmenting retinal blood vessels with deep neural networks. *IEEE Transactions on Medical Imaging*, 35(11) : 2369–2380, November 2016. *2 citations pages 36 et 38*
- [98] G. Litjens, T. Kooi, B.E. Bejnordi, A.A.A. Setio, F. Ciompi, M. Ghafoorian, J. Van Der Laak, B. Van Ginneken, and C.I. Sánchez. A survey on deep learning in medical image analysis. *Medical image analysis*, 42 : 60–88, December 2017. *Cité page 49*
- [99] L. Liu, P. Fieguth, G. Kuang, and H. Zha. Sorted random projections for robust texture classification. In *International Conference on Computer Vision*, pages 391–398, Barcelona, Spain, November 2011. *2 citations pages 91 et 95*
- [100] W.Y. Loh. Classification and regression trees. *Data Mining and Knowledge Discovery*, 1(1) : 14–23, January 2011. *2 citations pages 38 et 55*

- [101] C.P. Loizou, C.S. Pattichis, C.I. Christodoulou, R.S.H. Istepanian, M. Pantziaris, and A. Nicolaidis. Comparative evaluation of despeckle filtering in ultrasound imaging of the carotid artery. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 52(10) : 1653–1669, December 2005. *Cité page 35*
- [102] S.B. Lowen and M.C. Teich. Power-law shot noise. *IEEE Transactions on Information Theory*, 36(6) : 1302–1318, November 1990. *Cité page 37*
- [103] J. Lu, G. Wang, and P. Moulin. Image set classification using holistic multiple order statistics features and localized multi-kernel metric learning. In *International Conference on Computer Vision*, pages 329–336, Sydney, New South Wales, Australia, December 2013. *Cité page 80*
- [104] L. Maaten and G. Hinton. Visualizing data using t-sne. *Journal of Machine Learning Research*, 9 : 2579–2605, November 2008. *Cité page 42*
- [105] S. Mallat. *Une exploration des signaux en ondelettes*. Les Editions de l’Ecole Polytechnique, November 2000. *2 citations pages 88 et 161*
- [106] S. Mallat. Group invariant scattering. *Communications on Pure and Applied Mathematics*, 65(10) : 1331–1398, July 2012. *Cité page 78*
- [107] A. Mansour. *Probabilités et statistiques pour les ingénieurs : cours, exercices et programmation*. Hermes Science, November 2007. *10 citations pages 175, 176, 177, 178, 180, 181, 183, 184, 185 et 186*
- [108] A. Mansour, A.K. Barros, and N. Ohnishi. Comparison among three estimators for high order statistics. In *International Conference on Neural Information Processing*, pages 899–902, Kitakyushu, Japan, October 1998. *2 citations pages 186 et 187*
- [109] A. Mansour and C. Jutten. What should we say about the kurtosis? *IEEE Signal Processing Letters*, 6(12) : 321–322, December 1999. *2 citations pages 181 et 182*
- [110] A. Mansour, C. Jutten, and N. Ohnishi. Kurtosis : Definition and Properties. In *International Conference on Multisource-Multisensor : Data Fusion*, pages 40–46, Las Vegas, Nevada, United States, July 1998. *2 citations pages 181 et 182*
- [111] A. Martin and A. Mansour. High order statistic estimators for speech processing. In *International Conference on ITS Telecommunications*, pages 295–300, Brest, France, June 2005. *3 citations pages 80, 186 et 187*
- [112] M. Martín-Fernández and C. Alberola-Lopez. An approach for contour detection of human kidneys from ultrasound images using markov random fields and active contours. *Medical image analysis*, 9(1) : 1–23, February 2005. *Cité page 36*
- [113] MathWorks. Le deep learning à matlab. <https://fr.mathworks.com/discovery/deep-learning.html#withmatlab>. [Online : accessed on July 17, 2018]. *2 citations pages 47 et 48*
- [114] Nucleus Medical Media. Venous thrombus and embolus. <http://catalog.nucleusmedicalmedia.com/venous-thrombus-and-embolus/view-item?ItemID=4746>. [Online : accessed on July 17, 2018]. *Cité page 7*
- [115] J.M. Mendel. Tutorial on higher-order statistics (spectra) in signal processing and system theory : Theoretical results and some applications. *Proceedings of the IEEE*, 79(3) : 278–305, March 1991. *2 citations pages 80 et 180*
- [116] E. Mfoumou, J. Tripette, M. Blostein, and G. Cloutier. Time-dependent hardening of blood clots quantitatively measured in vivo with shear-wave ultrasound imaging in a rabbit model of venous thrombosis. *Thrombosis and Haemostasis*, 133(2) : 265–271, February 2014. *Cité page 31*
- [117] K.V. Mogatadakala, K.D. Donohue, C.W. Piccoli, and F. Forsberg. Detection of breast lesion regions in ultrasound images using wavelets and order statistics. *Medical physics*, 33(4) : 840–849, March 2006. *Cité page 37*
- [118] C. Molder, H. Thomas, and A. Quinquis. Classification des sédiments marins par analyse de texture. In *Congr  Reconnaissance des Formes et Intelligence Artificielle*, Angers, France, January 2002. *Cité page 37*

- [119] H. Monpeyssen, J. Tramalloni, S. Poirée, O. Hélénon, and J.M. Correas. L'élastographie de la thyroïde. *Journal de Radiologie Diagnostique et Interventionnelle*, 94(5) : 550–559, May 2013. *Cité page 35*
- [120] Y. Mori. *Électronique pour le traitement du signal : Notions de signal et de bruit*. Hermes Science, 2006. *3 citations pages 176, 187 et 188*
- [121] A. Nami. *Precision Imaging : Advanced 2D Imaging*. University Hospitals Cleveland Medical Center. Toshiba Medical Systems, 2010. <https://medical.toshiba.com/download/ul-wp-viamo-precision-imaging> [Online : accessed on May 31, 2016]. *Cité page 20*
- [122] A. Nasser, D. Hamad, and C. Nasr. K-means clustering algorithm in projected spaces. In *International Conference on Information Fusion*, pages 1–6, Florence, Italy, July 2006. *Cité page 50*
- [123] A.Y. Ng, M.I. Jordan, and Y. Weiss. On spectral clustering : Analysis and an algorithm. In *Advances in Neural Information Processing Systems*, pages 849–856, Cambridge, Massachusetts, USA, December 2002. *Cité page 54*
- [124] H.G. Nguyen, R. Fablet, and J.M. Boucher. Visual textures as realizations of multivariate log-Gaussian Cox processes. In *Conference on Computer Vision and Pattern Recognition*, pages 2945–2952, Colorado Springs, Colorado, USA, June 2011. *2 citations pages 91 et 95*
- [125] F. Nicolas, A. Arnold-Bos, I. Quidu, and B. Zerr. Markov-based approaches for ternary change detection between two high resolution synthetic aperture sonar tracks. In *OCEANS*, pages 1–9, Aberdeen, United Kingdom, June 2017. *Cité page 36*
- [126] D. Paoliello, T. Tan, and A. Mansour. Classification of electroencephalogram signals for human motor actions. In *World Congress of Biomechanics*, pages 1374–1377, Singapore, August 2010. *Cité page 37*
- [127] A. Parriaux and M. Arnould. *Géologie : bases pour l'ingénieur*. Presses polytechniques et universitaires romandes, September 2009. *Cité page 9*
- [128] B. Pesquet-Popescu and J. Pesquet. *Ondelettes et Applications*. Edition Techniques Ingénieur, August 2001. *8 citations pages 40, 159, 161, 163, 165, 167, 169 et 170*
- [129] P. Petrossians, S. Petignot, A. Benoit, and A. Beckers. *Echographie de la Thyroïde*. Lulu, April 2015. *Cité page 19*
- [130] C.G. Puntonet, A. Mansour, C. Bauer, and E. Lang. Separation of sources using simulated annealing and competitive learning. *Neurocomputing*, 49(1-4) : 39–60, December 2002. *Cité page 80*
- [131] J.R. Quinlan. Induction of decision trees. *Machine learning*, 1(1) : 81–106, March 1986. *2 citations pages 38 et 55*
- [132] A. Quinquis, I. Cornel, A. Mansour, and E. Radoi. *Le traitement du signal sous Matlab : pratique et applications*. Hermes Science, October 2007. *Cité page 179*
- [133] E. Radoi, F. Totir, A. Quinquis, and L. Anton. Superresolution imagery based SVM classification of radar targets. In *European Conference on Synthetic Aperture Radar*, pages 1–4, Dresde, Germany, May 2006. *Cité page 38*
- [134] Y. Rangsanseri and W. Prasongsook. Speckle reduction using wiener filtering in wavelet domain. In *International Conference on Neural Information Processing*, pages 792–795, Singapore, November 2002. *Cité page 35*
- [135] R. Riffiod. La circulation sanguine chez l'homme. <http://svtedu.free.fr/acad/tice/circulation/index.htm>, March 2011. [Online : accessed on July 17, 2018]. *Cité page 7*
- [136] K. Rousseeuw, E. Poisson Caillault, A. Lefebvre, and D. Hamad. Hybrid hidden Markov model for marine environment monitoring. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8(1) : 204–213, January 2015. *Cité page 36*
- [137] S.T. Roweis and L.K. Saul. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290(5500) : 2323–2326, December 2000. *Cité page 43*

- [138] J.M. Rubin, S.R. Aglyamov, T.W. Wakefield, M. O'Donnell, and S.Y. Emelianov. Clinical Application of Sonographic Elasticity Imaging for Aging of Deep Venous Thrombosis Preliminary Findings. *Journal of Ultrasound in Medicine*, 22(5) : 443–448, January 2003.
2 citations pages 31 et 32
- [139] M.S. Runge, A.M. Greganti, P.L. Masson, and J.S. Co. *Médecine interne de Netter*. Elsevier Health Sciences, April 2011.
Cité page 108
- [140] A. Samier-Guérin, A. Saraux, S. Gestin, D. Cornec, T. Marhadour, V. Devauchelle-Pensec, L. Bressollette, M. Nonent, and S. Jousse-Joulin. Can arfi elastometry of the salivary glands contribute to the diagnosis of sjögren's syndrome? *Joint Bone Spine*, 83(3) : 301–306, 2016.
3 citations pages 27, 35 et 108
- [141] J. Schmidhuber. Deep learning in neural networks : An overview. *Neural networks*, 61 : 85–117, January 2015.
Cité page 47
- [142] C. Schmitt. *L'élastographie ultrasonore dynamique vasculaire : une nouvelle modalité d'imagerie non-invasive pour la caractérisation mécanique de la thrombose veineuse*. PhD thesis, Université de Montréal, Montréal, Canada, April 2011.
3 citations pages 31, 32 et 33
- [143] B. Schölkopf, A. Smola, and K.R. Müller. Kernel principal component analysis. In *International Conference on Artificial Neural Networks*, pages 583–588, Lausanne, Switzerland, October 1997.
Cité page 43
- [144] E. Sellier, J. Labarere, M.A. Sevestre, J. Belmin, H. Thiel, P. Couturier, and J.L. Bosson. Risk factors for deep vein thrombosis in older patients : a multicenter study with systematic compression ultrasonography in postacute care facilities in France. *Journal of the American Geriatrics Society*, 56(2) : 224–230, February 2008.
Cité page 5
- [145] P.M. Shankar. A general statistical model for ultrasonic backscattering from tissues. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 47(3) : 727–736, May 2000.
Cité page 37
- [146] P.M. Shankar, V.A. Dumane, T. George, C.W. Piccoli, J.M. Reid, F. Forsberg, and B.B. Goldberg. Classification of breast masses in ultrasonic B scans using Nakagami and K distributions. *Physics in Medicine and Biology*, 48(14) : 2229–2240, July 2003.
2 citations pages 37 et 38
- [147] G. Sharma, S. ul Hussain, and F. Jurie. Local higher-order statistics (LHS) for texture categorization and facial analysis. In *European Conference on Computer Vision*, pages 1–12, Florence, Italy, October 2012.
Cité page 80
- [148] H. Shin, H.R. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, J. Yao, D. Mollura, and R.M. Summers. Deep Convolutional Neural Networks for Computer-Aided Detection : CNN Architectures, Dataset Characteristics and Transfer Learning. *IEEE Transactions on Medical Imaging*, 35(5) : 1285–1298, May 2016.
2 citations pages 38 et 49
- [149] S. Siebers, B. Geier, D. Muth-Werthmann, A. Mumme, T. von Rothenburg, S. Philippou, and H. Ermert. Staging of venous thrombosis using ultrasound elastography. In *IEEE Symposium on Ultrasonics*, pages 1891–1894, Honolulu, Hawaii, USA, October 2003. 2 citations pages 31 et 32
- [150] S. Siebers, B. Geier, U. Scheipers, M. Vogt, A. Mumme, and H. Ermert. Classification of venous thrombosis combining ultrasound elastography and tissue characterization. In *IEEE Ultrasonics Symposium*, pages 1761–1764, Montreal, Quebec, Canada, August 2004.
3 citations pages 30, 31 et 32
- [151] L. Sifre and S. Mallat. Rotation, scaling and deformation invariant scattering for texture discrimination. In *Conference on Computer Vision and Pattern Recognition*, pages 1233–1240, Portland, Oregon, USA, June 2013.
6 citations pages 72, 77, 79, 91, 95 et 104
- [152] M. Singh, S. Singh, and S. Gupta. An information fusion based method for liver classification using texture analysis of ultrasound images. *Information Fusion*, 19 : 91–96, September 2014.
2 citations pages 37 et 38

- [153] K. Skerl, S. Vinnicombe, S. McKenna, K. Thomson, and A. Evans. First step for computer assisted evaluation of qualitative supersonic shear wave elastography characteristics in breast tissue. In *IEEE International Symposium on Biomedical Imaging*, pages 481–484, Prague, Czech Republic, April 2016. *Cité page 35*
- [154] J.H. Song, S.S. Venkatesh, E.F. Conant, T.W. Cary, P.H. Arger, and C.M. Sehgal. Artificial neural network to aid differentiation of malignant and benign breast masses by ultrasound imaging. In *Medical Imaging : Ultrasonic Imaging and Signal Processing*, San Diego, California, United States, February 2005. *2 citations pages 37 et 38*
- [155] H. Stiegler, H. Arbogast, S. Nees, A. Halder, A. Grau, and H. Riess. Thrombectomy, lysis, or heparin treatment : concurrent therapies of deep vein thrombosis : therapy and experimental studies. *Seminars in Thrombosis and Hemostasis*, 15(3) : 250–258, July 1989. *Cité page 30*
- [156] M. Stone. Cross-validatory choice and assessment of statistical predictions. *Journal of the Royal Statistical Society. Series B (Methodological)*, 36(2) : 111–147, December 1974. *Cité page 92*
- [157] Toshiba Medical Systems. *Aplio Brochure : Radiology and Shared Service Transducers*, 2016. <https://medical.toshiba.com/download/ul-br-aplio-transducers.pdf> [Online : accessed on October 16, 2017]. *Cité page 18*
- [158] G. Tartare, D. Hamad, M. Azahaf, P. Puech, and N. Betrouni. Spectral clustering applied for dynamic contrast-enhanced MR analysis of time-intensity curves. *Computerized Medical Imaging and Graphics*, 38(8) : 702–713, December 2014. *Cité page 36*
- [159] J.B. Tenenbaum, V. De Silva, and J.C. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, 290(5500) : 2319–2323, December 2000. *Cité page 43*
- [160] Toshiba Medical Systems Corporation. *System Brochure : Aplio 500, picture perfect ultrasound*, 2012. http://www.toshibamedicalsystems.com/products/us/pdf/brochure_aplio500.pdf [Online : accessed on July 13, 2016]. *2 citations pages 15 et 16*
- [161] N. Valeyrie, Y. Pailhas, C. Capus, and Y. Petillot. Texture recognition in synthetic aperture sonar images with scattering operators. In *International Conference and Exhibition on Underwater Acoustic Measurements : Technologies and Results*, Kos Island, Greece, June 2011. *3 citations pages 55, 72 et 79*
- [162] A. Vedaldi and K. Lenc. Matconvnet : Convolutional neural networks for matlab. In *ACM Multimedia Conference*, pages 689–692, Brisbane, Australia, October 2015. *Cité page 48*
- [163] U. Von Luxburg. A tutorial on spectral clustering. *Statistics and Computing*, 17(4) : 395–416, December 2007. *2 citations pages 39 et 51*
- [164] C.M. Wu and Y.C. Chen. Statistical feature matrix for texture analysis. *CVGIP : Graphical Models and Image Processing*, 54(5) : 407–419, September 1992. *Cité page 37*
- [165] L. Zelnik-Manor and P. Perona. Self-tuning spectral clustering. In *Conference on Neural Information Processing Systems*, pages 1601–1608, Vancouver, British Columbia, Canada, December 2004. *Cité page 195*
- [166] K. Zuiderveld. Contrast limited adaptive histogram equalization. In *Graphics Gems IV*, pages 474–485. Academic Press, San Diego, CA, USA, May 1994. *2 citations pages 36 et 64*

Titre : Identification et caractérisation du thrombus veineux par imagerie échographique mode B couplée à l'élastographie

Mots clés : Imagerie acoustique, Thrombose veineuse profonde, Ondelettes, Statistiques d'ordre supérieur, Classification

Résumé : La maladie veineuse thromboembolique (MVTE) est un problème de santé publique (plus de 100 000 cas par an en France). Elle regroupe deux entités cliniques : la thrombose veineuse profonde (TVP) des membres inférieurs et l'embolie pulmonaire (EP). La TVP correspond à la formation inadaptée d'un thrombus veineux (appelé aussi caillot sanguin) dans les veines profondes (poplitées, fémorales, iliaques). Un thrombus est principalement constitué de globules rouges et de plaquettes dans un réseau de fibrine. La complication majeure d'une TVP est la survenue d'une EP, c'est-à-dire que le thrombus s'est détaché de la paroi veineuse, ou s'est fragmenté, et est entraîné par la circulation sanguine jusqu'à une artère pulmonaire. Cette complication a un taux de mortalité assez élevée autour de 10 000 à 20 000 cas mortel par an en France.

La survenue d'une TVP est multifactorielle

associant des facteurs génétiques et acquis pouvant être répartis en trois catégories : la stase veineuse, l'altération de la paroi d'une veine et une hypercoagulabilité. En analysant la structure du thrombus, notre projet vise à identifier le facteur principal responsable de la TVP et à évaluer le risque d'EP. Pour caractériser sa structure, nous disposons de deux modes d'imagerie acoustique : l'échographie et l'élastographie (carte de dureté). Nous proposons d'extraire des descripteurs de ces images acoustiques par deux approches, l'une basée sur les ondelettes (le *scattering operator*) et l'autre sur les statistiques d'ordre supérieur (les multicorrélations). Ces descripteurs sont ensuite analysés par diverses techniques de classification (analyse en composantes principales, *k*-moyennes, classification spectrale) pour retrouver la cause principale des TVP ou la présence d'EP.

Title: Venous Thrombus identification and characterization using ultrasonography and elastography

Keywords: Ultrasound image, Deep venous thrombosis, Wavelet, High order statistics, Classification

Abstract: Venous thromboembolism (VTE) is an important public health issue (over 100 00 individuals in France per year). VTE is a combination of a deep venous thrombosis (DVT) and a Pulmonary Embolism (PE). DVT is an inappropriate formation of a thrombus (also called blood clot) in one of the deep veins of the body, usually in the leg (popliteal, femoral, iliac). There are mainly three components in a thrombus: platelets, red blood cells and a mesh of fibrins. The main complication of a DVT is a Pulmonary Embolism (PE) which occurs when a thrombus breaks loose and travels to the lungs. PE affects an estimated 10 000–20 000 individuals just in France per year.

Three physiopathological mechanisms can contribute, isolated or combined, to the development

of a DVT: venous stasis, endothelial injury and hypercoagulability. Our project is aiming to relate the thrombus structure, its main triggering factor and the risk of a PE. To characterise the thrombus structure, we are collecting ultrasonography (echogenicity) and elastography (stiffness) of human thrombus. We propose to extract features from these to kind of ultrasound images using two approaches: one based on wavelets (the *scattering operator*) and another based on high order statistics (multicorrelations). Then, the obtained features are analysed using several classification technics (principal component analysis, *k*-means, spectral clustering) to find the main cause of the DVT or the presence of PE.