

Traitement et analyse de signaux sonores physiologiques : application à la phonoentérographie

Radu Ranta

▶ **To cite this version:**

Radu Ranta. Traitement et analyse de signaux sonores physiologiques : application à la phonoentérographie. Autre. Institut National Polytechnique de Lorraine, 2003. Français. NNT : 2003INPL100N . tel-01749833

HAL Id: tel-01749833
https://hal.univ-lorraine.fr/tel-01749833

Submitted on 29 Mar 2018

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.

AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact : ddoc-theses-contact@univ-lorraine.fr

LIENS

Code de la Propriété Intellectuelle. articles L 122. 4
Code de la Propriété Intellectuelle. articles L 335.2- L 335.10
http://www.cfcopies.com/V2/leg/leg_droi.php
http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm

136035175 5
03INPL100N
[TT] 2003 RANTA R.

Ecole doctorale IAEM Lorraine
DFD Automatique et Production Automatisée

Institut National Polytechnique de Lorraine

Traitement et analyse de signaux sonores physiologiques

Application à la phonoentérographie

Service Commun de la Documentation
INPL
Nancy-Brabois

THÈSE

présentée et soutenue publiquement le 9 décembre 2003

pour l'obtention du

Doctorat de l'Institut National Polytechnique de Lorraine

Spécialité Automatique et Traitement du Signal

par

Radu RANTA

Composition du jury

Rapporteurs :	Régine Le Bouquin-Jeannès	LTSI Rennes
	Gérard Gimenez	CERMEP – CREATIS Lyon
Examinateurs :	Valérie Louis-Dorr	CRAN Nancy
	Christian Heinrich	LSIIT Strasbourg
	François Guillemin	CAV – CRAN Nancy
	Didier Wolf	CRAN Nancy
Invité :	Jacques Duchêne	LM2S Troyes

Centre de Recherche en Automatique de Nancy CNRS – UMR 7039

Mis en page avec la classe thloria.

Remerciements

Cette étude a été menée au Centre de Recherche en Automatique de Nancy, dans le thème Ingénierie pour la Santé dirigée par Monsieur le Professeur Didier Wolf, qui a co-encadré mes recherches avec Madame Valérie Louis-Dorr, Maître de conférences à l'INPL.

Je tiens tout d'abord à remercier mes encadrants pour m'avoir accueilli dans l'équipe, pour leur suivi et pour leurs conseils, ainsi que pour l'atmosphère à la fois travailleuse et agréable qu'ils ont su créer et entretenir.

Je remercie Monsieur le Professeur François Guillemin de la Faculté de Médecine de Nancy, qui a initié et suivi avec intérêt ce travail et qui m'a fait également l'honneur de présider mon jury de thèse.

Je remercie aussi Monsieur Christian Heinrich, Maître de conférences à l'Université Louis Pasteur de Strasbourg, qui a participé à l'encadrement de mes travaux de DEA et au jury de ma thèse, et avec lequel j'ai gardé une fructueuse collaboration.

Je tiens à remercier les membres du jury qui m'ont fait l'honneur de participer à l'examen de ce travail.

Madame le Professeur Régine Le Bouquin-Jeannès, de l'Université de Rennes 1, pour l'attention qu'elle a accordé à ce travail et pour sa présence en tant que rapporteur dans ce jury.

Monsieur le Professeur Gérard Gimenez, de l'Institut National de Sciences Appliquées (INSA) de Lyon, pour avoir accepté d'examiner ce travail en tant que rapporteur.

Monsieur le Professeur Jacques Duchêne, de l'Université de Technologie de Troyes, pour sa participation à ce jury.

Je remercie Monsieur Gilles Mourot, Ingénieur de recherche au CRAN, et Monsieur Thierry Loyau, chercheur à l'INRS, pour les conseils qu'ils m'ont apportés au cours de mes travaux, ainsi que tous les membres du laboratoire pour leur soutien.

Enfin, je tiens à remercier mes collègues pour leur intérêt, leur aide et surtout leur amitié.

Merci à Dana, merci à ma famille.

Table des matières

Introduction 1

Chapitre 1 Sons abdominaux. Contexte et problématique 5

1.1 Contexte et objectifs médicaux 6

1.1.1 Anatomie et physiologie 6

1.1.2 Méthodes d'investigation 10

Les signaux abdominaux 10

Les sons abdominaux 11

1.1.3 Conclusion 14

1.2 Problématique scientifique 16

1.2.1 Signal et instrumentation 16

1.2.2 Pré-traitement 17

1.2.3 Analyse des signaux et des événements 19

1.3 Conclusion 20

Chapitre 2 Acquisition des sons : signal et instrumentation 23

2.1 Signal 23

2.2 Protocoles de mesure 25

2.2.1 Proposition de protocole 26

2.3 Chaîne d'instrumentation 27

2.3.1 Capteurs 28

2.3.2 Conditionnement et numérisation 28

2.3.3 Chaîne d'instrumentation multi-voies 29

2.4 Étude de la réponse fréquentielle 29

2.4.1 Méthode 29

2.4.2 Protocole 32

Instrumentation et conditions de test 32

Déroulement du protocole 36

2.4.3 Résultats 37

Réponses fréquentielles 37

Influence de la pression 38
Discussion 40
2.5 Conclusion 41

Chapitre 3 Détection et débruitage par ondelettes **43**
3.1 Analyse temps-fréquence. Ondelettes 44
3.1.1 Décompositions atomiques 46
3.1.2 Bases d'ondelettes orthogonales 48
3.2 Détection, estimation et débruitage par ondelettes 50
3.2.1 Stratégies de seuillage 50
3.2.2 Calcul du seuil 51
Méthodes directes 51
Méthodes itératives 52
3.3 Algorithme de débruitage de Hadjileontiadis *et al.* 54
3.4 Interprétation de point-fixe 56
3.4.1 Préliminaires 56
3.4.2 Seuillage 56
3.4.3 Interprétation de point fixe 57
3.4.4 Conditions générales d'application 57
3.5 Modélisation probabiliste 59
3.5.1 Modèles gaussiens généralisés 61
3.5.2 Optimisation 62
3.6 Segmentation 64
3.7 Validation 66
3.7.1 Signaux considérés 66
3.7.2 Critères de performance 68
3.7.3 Algorithmes testés 71
3.7.4 Base d'ondelettes 71
3.7.5 Résultats 72
3.8 Conclusion 80

Chapitre 4 Segmentation et localisation **85**
4.1 Préliminaires : extraction des caractéristiques physiques 85
4.1.1 Ondelettes et caractéristiques physiques 86
4.1.2 Indices d'activité 93
4.2 Segmentation : élimination des artéfacts 94
4.2.1 Algorithme d'élimination des artéfacts 95

4.2.2 Résultats, exemples 97
4.3 Localisation 101
4.3.1 Méthodes de localisation 101
4.3.2 Améliorations 103
4.3.3 Résultats, exemples 109
4.4 Conclusion 111

Chapitre 5 Analyse des phonoentérogrammes **113**
5.1 Analyse en composantes principales 114
5.1.1 Espaces de caractéristiques 114
Indices d'activité 116
Caractéristiques physiques 121
5.1.2 Analyse des données selon l'ACP 123
Séquences de phonoentérogramme 125
Sons abdominaux individuels 131
5.2 Classification non-supervisée 132
5.2.1 Méthodes de classification non-supervisée 133
5.2.2 Hypothèses de travail et algorithme 138
Analyse discriminante 139
Algorithme de classification non-supervisée 140
5.2.3 Résultats 144
5.3 Conclusion, perspectives 148

Conclusion générale. Perspectives **149**

Annexes **153**

Annexe A Ondelettes testées **155**

Annexe B Validation algorithmes de débruitage. Résultats détaillés **157**

Bibliographie **165**

Table des figures

1.1 Anatomie générale du tube digestif – partie abdominale 7
1.2 Étapes de traitement 17

2.1 Exemple de signal sonore abdominal 24
2.2 Exemple de spectre d'un signal sonore abdominal 24
2.3 Exemple de spectre du bruit 24
2.4 Placement des capteurs 26
2.5 Exemple de réponses fréquentielles des stéthoscopes 27
2.6 Chaîne d'instrumentation 28
2.7 Modèle d'identification simplifié 30
2.8 Protocole de mesure – étapes de calibrage 33
2.9 Spectre de la source sonore de test 34
2.10 Estimation de la stationnarité de la source de test 36
2.11 Protocole de mesure – étapes d'estimation 36
2.12 Dispositif pour l'étude de l'influence de la pression d'appui 37
2.13 Protocole de mesure – pression d'appui 37
2.14 Réponse fréquentielle du microphone utilisé 38
2.15 Réponses fréquentielles des capteurs 38
2.16 Réponses fréquentielles en fonction de la pression (capteur 1) 39
2.17 Réponses fréquentielles en fonction de la pression (capteur 2) 39

3.1 Représentation temporelle d'un signal 45
3.2 Représentation fréquentielle d'un signal 45
3.3 Représentation temps-fréquence d'un signal 46
3.4 Atome temps-fréquence 47
3.5 Algorithme de décomposition en ondelettes 49
3.6 Algorithme itératif de débruitage par ondelettes 53
3.7 Histogramme des coefficients d'ondelette 55
3.8 Fonction $f(S)$ constante par morceaux 58
3.9 Condition d'existence du point fixe 60
3.10 Condition d'existence du point fixe – limitations 61
3.11 Exemples de gaussiennes généralisées 63
3.12 Variation de F_{am}^{j} 63
3.13 Variation de la constante K_c 63
3.14 Variation de F_{am}^{j} 63
3.15 Segmentation 65
3.16 Signaux de test artificiels de Donoho et Johnstone 67
3.17 Signaux de test simulés 68
3.18 Spectres du bruit blanc et du bruit coloré simulés 69

3.19 Histogrammes des bruits additifs 70
3.20 Exemple de débruitage inadéquat 74
3.21 Histogrammme particulier des coefficients d'ondelette (*Bumps*) 74
3.22 Résultats débruitage, bruit blanc laplacien 78
3.23 Résultats débruitage, bruit blanc laplacien 79
3.24 Spectre du signal simulé à 5 événements par minute 81
3.25 Seuils des différents algorithmes de débruitage, par échelle 81
3.26 Résultats débruitage, signal réel 1 83
3.27 Résultats débruitage, signal réel 2 84

4.1 Spectres des ondelettes Daubechies no. 9 87
4.2 Pavés temps-fréquence et ondelettes correspondantes 88
4.3 Segmentation et pavés temps-fréquence 90
4.4 Enveloppe spectrale 91
4.5 Enveloppe par ondelettes 92
4.6 Évolution de l'indice N_m – nombre d'événements 97
4.7 Évolution des indices D_m (durée totale) et E_m (énergie totale) 98
4.8 Étapes de re-segmentation 100
4.9 Espace des solutions possibles pour la localisation 106
4.10 Erreur de modèle de localisation 107
4.11 Variation spatiale de l'activité abdominale au cours du temps (modèle isotrope) . 108
4.12 Variation spatiale de l'activité abdominale au cours du temps (modèle empirique) 109
4.13 Évolution de l'indice N_m par voie, durant 168 minutes 110

5.1 Exemple de variations des indices d'activité 118
5.2 ACP indices d'activité – cercles des corrélations 120
5.3 Répartition des séquences dans l'espace engendré par les 3 premières composantes principales 122
5.4 ACP caractéristiques physiques – cercles des corrélations 124
5.5 Répartition des événements dans l'espace engendré par les 3 premières composantes principales 124
5.6 Projections des voies v_1 à v_6 sur les plans principaux 126
5.7 Répartition 3D des séquences des voies v_3 et v_4 126
5.8 Projections des voies v_1 à v_6 sur les plan $c_2 - c_3$ pour les trois patients 127
5.9 Projections des séquences s_1 à s_{12} sur les plans principaux 127
5.10 Évolution temporelle des séquences selon les axes principaux 128
5.11 Évolution des séquences s_1–s_{12} selon l'axe c_3 pour les trois patients 128
5.12 Projections des patients p_1 à p_3 sur les plans principaux 128
5.13 Comparaison des évolutions de c_2 pour la voie v_4 129
5.14 Projections des phonoentérogrammes normaux p_1–p_3 et du signal pathologique sur les plans principaux 130
5.15 Projections des voies v_1 à v_6 sur les plans principaux (espace des sons) 132
5.16 Projections des sons individuels normaux et des sons pathologiques sur les plans principaux 132
5.17 Exemple de classification hiérarchique 136
5.18 Organigramme de l'algorithme de classification non-supervisée 141
5.19 Histogramme des coefficients d'inconsistance 142
5.20 Histogramme des séquences de phonoentérogrammes après algorithme hiérarchique 143

5.21 Comparaison classifications en fonction du seuil de fusion 145
5.22 Répartition par classes des séquences des patients 146
5.23 Répartition et évolution en classes des séquences, voie par voie 147

A.1 Fonction d'échelle et ondelette Daubechies no.9 155
A.2 Fonction d'échelle et ondelette Symlets no.9 155

B.1 Critère de débruitage EQM pour le un bruit additif blanc. Détails 158
B.2 Critère de débruitage EQM pour le un bruit additif coloré. Détails 158
B.3 Critère de débruitage EQM pour le un bruit additif blanc. Détails 159
B.4 Critère de débruitage EQM pour le un bruit additif coloré. Détails 159
B.5 Critère de débruitage ES pour le un bruit additif blanc. Détails 160
B.6 Critère de débruitage ES pour le un bruit additif coloré. Détails 160
B.7 Critère de débruitage ES pour le un bruit additif blanc. Détails 161
B.8 Critère de débruitage ES pour le un bruit additif coloré. Détails 161
B.9 Critère de débruitage ED pour le un bruit additif blanc. Détails 162
B.10 Critère de débruitage ED pour le un bruit additif coloré. Détails 162
B.11 Critère de débruitage ED pour le un bruit additif blanc. Détails 163
B.12 Critère de débruitage ED pour le un bruit additif coloré. Détails 163

Introduction

L'ingénierie biomédicale suscite un intérêt grandissant depuis plusieurs années. Pluridisciplinaire, elle progresse grâce aux avancées dans des domaines aussi divers que la génétique ou le traitement du signal. C'est dans ce dernier domaine que notre recherche s'inscrit : le but de ce travail est de développer un outil de recherche et d'aide au diagnostic médical fondé sur l'analyse des sons physiologiques, et plus particulièrement des sons abdominaux. C'est un travail méthodologique de traitement de signal qui a comme objectif médical fondamental sur le long terme la construction d'un outil pour l'étude de la physiologie normale ou pathologique des organes abdominaux.

Un des moyens d'investigation physiologique les plus anciens et toujours très utilisé en routine clinique est l'auscultation. C'est une méthode peu onéreuse, car l'instrumentation est simple (stéthoscope) et son utilité est largement reconnue surtout pour les sons cardiaques et pulmonaires, mais aussi pour les sons abdominaux et fœtaux. Relativement peu étudiée pour les sons physiologiques abdominaux, elle présente des potentialités importantes car ces sons sont riches d'informations tant pour la compréhension de l'action mécanique qui les génère que pour l'interprétation physiologique que l'on peut en faire [15, 18, 48, 49]. Malgré l'intérêt suscité par les sons abdominaux depuis relativement longtemps [18], les conclusions médicales restent peu nombreuses. Cette situation s'explique par la difficulté de l'interprétation des sons, due à leur caractère irrégulier, à leur apparition aléatoire, à la variabilité entre différents patients et enfin, à la nécessité d'une auscultation de longue durée. Tous ces éléments rendent l'interprétation difficile et subjective pour les experts médicaux.

D'où l'intérêt de développer un outil automatisé de diagnostic clinique et de recherche médicale fondé sur le traitement systématique des signaux sonores physiologiques. Un tel outil permettrait également de procéder à une surveillance sur de longues durées sans mobiliser un médecin. Notons que l'auscultation abdominale automatisée, apparue au milieu du $XX^{ième}$ siècle, suscite un regain d'intérêt depuis une décennie, en raison d'avancées technologiques permettant un traitement automatisé des données [55, 116].

On estime que les bruits intestinaux traduisent l'activité motrice des différents segments du tract gastro-intestinal : lors de la digestion, le bol alimentaire est transformé et transporté le long du tube digestif, et le contenu liquide et gazeux intra-viscéral provoque des bruits. A la différence des examens d'imagerie médicale, qui donnent une information plutôt anatomique et statique de l'abdomen, l'interprétation des sons offre des informations physiologiques complémentaires. Outre l'étude du fonctionnement dans des conditions normales, la méthode peut être appliquée en médecine clinique, notamment pour l'aide au diagnostic des différentes maladies fonctionnelles, pour la surveillance post-chirurgicale, ainsi que dans la recherche pharmacologique (étude de l'effet de certains médicaments sur l'activité gastro-intestinale).

L'objectif scientifique qui découle des précédentes considérations est le développement d'un dispositif automatisé d'auscultation, de traitement et d'analyse des sons abdominaux, utilisé par la suite comme un outil de recherche et/ou comme une aide au diagnostic médical. Cet objectif comporte plusieurs volets, qui sont d'ailleurs communs à tous les outils d'aide à la décision [40] :

1. **L'acquisition** des signaux. L'instrumentation doit permettre un enregistrement fidèle des signaux d'intérêt, tout en minimisant l'incidence des signaux parasites. Elle doit également être robuste face aux éventuelles variations dans les conditions d'utilisation clinique et faciliter la validation médicale ;
2. **Le pré-traitement** du signal recueilli. L'objectif de cette étape est la transformation du signal brut en un signal interprétable, par débruitage, segmentation et localisation ;
3. **La caractérisation** des signaux enregistrés. Cette étape consiste en le choix et/ou la définition de caractéristiques permettant une description des signaux informatifs, éventuellement sur plusieurs niveaux de détails (caractérisation **globale** et **locale**) ;
4. **L'analyse des données et l'interprétation des résultats**. Les informations obtenues lors des étapes précédentes doivent être intégrées dans des algorithmes permettant de faire la distinction entre les différents types de signaux. La classification obtenue doit avoir une signification (séparer des signaux normaux des signaux pathologiques, par exemple) et doit pouvoir être interprétée par un médecin.

Cette thèse s'intéresse aux différents aspects énumérés ci-dessus. L'objectif de nos recherches est la construction d'une "boîte à outils" capable de traiter les phonoentérogrammes et de fournir aux médecins des données interprétables du point de vue clinique : nous cherchons à trouver des méthodes pour distinguer différents modes de fonctionnement, et non pas à interpréter ces différences.

Avant de nous lancer dans la conception des outils nécessaires pour atteindre cet objectif, il nous semble indispensable de consacrer un premier chapitre à l'étude de la physiologie abdominale et surtout de ses manifestations sonores. L'état de l'art que nous présentons dans ce chapitre place nos travaux dans le contexte médical de la phonoentérographie et énumère les méthodes de traitement et d'analyse décrites dans la littérature. Cette présentation permet de définir la problématique scientifique qui constituera la suite de nos travaux.

Le deuxième chapitre est consacré à l'acquisition des phonoentérogrammes. Nous présentons d'abord nos constats expérimentaux sur les signaux à traiter et nous définissons un protocole de mesure adapté à la problématique médicale, ce qui permet de concevoir une chaîne d'instrumentation dédiée. Nous étudions ensuite cette chaîne, afin de s'assurer que ses caractéristiques sont adaptées aux signaux de phonoentérographie et surtout que les grandeurs d'influence comme la pression d'appui sur l'abdomen ne changent pas de manière importante la réponse fréquentielle.

Le troisième chapitre de ce mémoire présente nos travaux sur la détection, le débruitage et la segmentation des phonoentérogrammes. Nous proposons un cadre de travail (représentation par ondelettes) dans lequel nous développons un nouvel algorithme de détection et débruitage applicable aux signaux non-stationnaires en général. Cet algorithme est ensuite testé sur un ensemble de signaux simulés et réels et ses performances sont comparées aux autres algorithmes décrits dans la littérature.

Le quatrième chapitre propose une approche pragmatique qui introduit des connaissances *a priori* sur les caractéristiques physiques des sons abdominaux. Nous proposons une méthode pour extraire ces caractéristiques à partir de la décomposition en ondelettes de chaque son et nous les comparons aux connaissances *a priori* afin de détecter et d'éliminèr les artéfacts superposés au signal informatif. Les caractéristiques physiques sont ensuite utilisées pour tester différents modèles de propagation des sons à l'intérieur de l'abdomen et pour choisir une méthode de localisation.

Le cinquième et dernier chapitre aborde le domaine de l'analyse et de l'interprétation des phonoentérogrammes. Nous proposons une analyse sur deux niveaux : un niveau global, qui

propose une caractérisation et une classification des différents types d'activité abdominale à l'aide de grandeurs moyennes évoluant dans le temps (indices d'activité), et un niveau local, qui vise à caractériser et classer les sons individuels. Ce chapitre, après une brève description des méthodes d'analyse de données comme l'analyse en composantes principales, la classification non-supervisée et l'analyse discriminante, est essentiellement dédié aux exemples de classification et d'interprétation possible des résultats.



Sons abdominaux. Contexte et problématique

Le but principal de l'ingénierie biomédicale est la conception d'outils d'investigation clinique et paraclinique, qui permettent la réalisation de deux grandes classes d'examens : les examens morphologiques (anatomiques) et les examens fonctionnels (physiologiques). Par leur nature, les examens anatomiques relèvent principalement du domaine de l'imagerie médicale, qui est le sujet de prédilection de l'ingénierie biomédicale, car la quantité d'informations apportée est grande et les données assez facilement interprétables. Pour compléter l'information morphologique fourni par les examens anatomiques, il faut recourir aux examens fonctionnels, dont une grande partie des techniques utilisées s'apparente aussi à l'imagerie. L'inconvénient d'un nombre important de techniques d'imagerie fonctionnelle est qu'elles sont invasives : elles nécessitent par exemple l'ingestion ou l'injection de substances de contraste, l'irradiation du patient ou l'introduction dans le corps d'appareils de mesure. C'est le cas notamment des investigations du tractus digestif par endoscopie ou par manométrie.

Il existe des techniques non invasives, comme l'électrocardiographie ou l'électroencéphalographie, mais elles n'ont pas vraiment d'équivalent dans le domaine de la gastroentérologie. L'électrogastrographie (EGG) [4, 22, 23], ainsi que l'analyse des vibrations de surface (SVA - *surface vibration analysis*) [17, 33] ne sont pas utilisées en routine clinique. En revanche, l'auscultation abdominale est un des temps de l'examen clinique. Elle est simple, non invasive et informative dans la démarche diagnostique. Cependant, l'auscultation abdominale reste délicate, pour plusieurs raisons :

- le signal n'est pas périodique ou pseudo-périodique, à la différence des sons cardiaques ou pulmonaires. Les sons potentiellement informatifs peuvent survenir à des instants irréguliers et espacés dans le temps ;
- la variabilité des sons rend leur interprétation difficile et souvent subjective ;
- l'information est répartie sur une longue durée, le cumul des événements étant plus informatif que les événements pris individuellement.

Les sons physiologiques sont porteurs d'informations, et le développement d'un outil d'investigation automatique fondé sur leur interprétation présente un grand intérêt.

Ce chapitre présente l'état de l'art et de la connaissance sur les méthodes et les résultats des investigations physiologiques abdominales, et en particulier sur l'étude des sons abdominaux : leur rapport avec l'anatomie et la physiologie de l'appareil digestif, l'information que l'on es-

time contenue et que l'on recherche, les méthodes utilisées pour l'extraire et leur interprétation physiologique.

A partir de ces informations, la problématique liée aux méthodes d'acquisition, traitement et interprétation des données est présentée dans la deuxième partie. Le développement des outils nécessaires au traitement et à l'analyse de ce type de signaux constitue l'essentiel de ce travail de recherche.

1.1 Contexte et objectifs médicaux

Dans cette partie, nous décrivons brièvement l'anatomie et la physiologie du tractus digestif et les mécanismes qui peuvent être considérés comme responsables de la génération des sons dans différentes conditions physiologiques. Cette étude doit nous permettre d'introduire les protocoles d'acquisition et de traitement utilisés par la suite.

1.1.1 Anatomie et physiologie

Le tube digestif, après la bouche et le pharynx, se poursuit par l'oesophage, qui traverse le cou et le thorax et se termine au cardia, orifice supérieur de l'estomac. La partie abdominale de l'appareil digestif commence par l'estomac, divisé en une partie haute, appelée la grosse tubérosité (poche à air), suivie par une partie intermédiaire, le corps. En bas se trouve l'antre, puis le pylore, par lequel l'estomac communique avec l'intestin. L'intestin comprend d'abord l'intestin grêle, dont la longueur varie de 4 à 7 mètres ; il débute par le duodénum, dont la première partie, renflée, est le bulbe duodénal. Au duodénum fait suite le jéjunum, qui représente environ 2/5 de l'intestin grêle, tandis que les 3/5 suivants s'appellent iléon. L'iléon se jette dans le gros intestin (ou côlon). Le gros intestin comprend, à droite, le cæcum, où s'abouche un diverticule, l'appendice. Au-dessus de l'appendice débouche l'iléon, protégé d'un reflux du contenu colique par un système de valve anti-retour (valve iléo-cæcale). Le cæcum est prolongé en haut par le côlon ascendant jusqu'à l'angle droit, puis par une partie horizontale, le côlon transverse, qui traverse la partie haute de l'abdomen, dessine l'angle gauche, et se continue par une partie gauche verticale, le côlon descendant. A celui-ci fait suite le côlon sigmoïde, qui décrit une anse et aboutit au rectum, lequel se termine par l'anus (voir figure 1.1) [113, 114].

La physiologie de l'appareil digestif est complexe. Globalement, on décrit quatre fonctions principales :

- une fonction sécrétoire, qui consiste en la production et l'évacuation dans le tube digestif des substances qui participent à la transformation des aliments en composants plus simples, assimilables dans le sang ;
- une fonction motrice, qui comporte le stockage des aliments, leur mélange aux substances sécrétées, ainsi que leur propulsion le long du tube digestif, de l'œsophage vers le rectum ;
- une fonction d'absorption, qui permet le passage des nutriments du tube digestif dans le sang ;
- une fonction endocrinienne, qui consiste en la production des substances à rôle régulateur de l'activité gastro-intestinale et de l'absorption des éléments nutritifs dans l'organisme.

L'hypothèse fondamentale de nos travaux est que les sons abdominaux sont une signature de la motricité des différents segments du tube digestif. On considère que l'énorme majorité des bruits, qui sont de type impulsionnel, sont provoqués par les transformations mécaniques et les déplacements du bol alimentaire dus aux contractions musculaires des parois du tube digestif. Les appuis bibliographiques de cette hypothèse sont présentés dans les sections suivantes, qui traitent du fonctionnement normal et pathologique de l'appareil digestif.

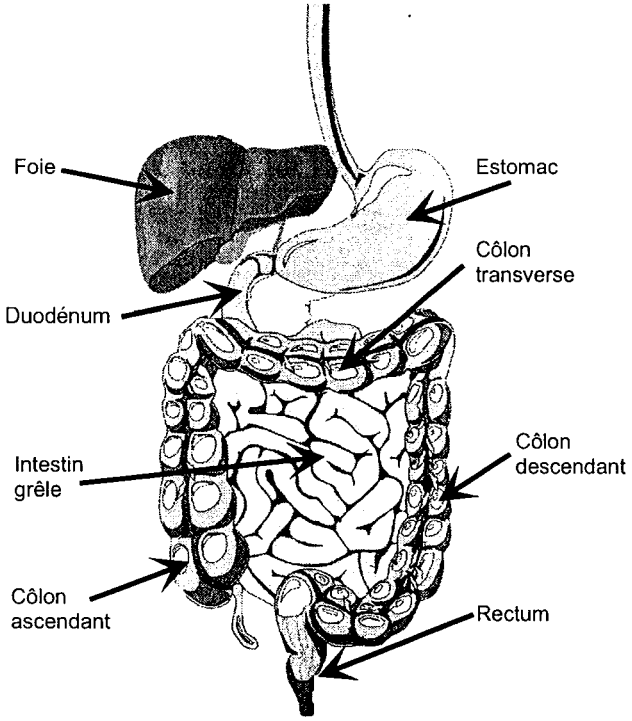


FIG. 1.1 – Anatomie générale du tube digestif – partie abdominale.

Avant d'aborder les manifestations sonores de la motricité gastro-intestinale, nous considérons utile une présentation de cette fonction du tube digestif, telle qu'elle est observée grâce à d'autres méthodes d'investigation, comme l'électrogastrographie et la manométrie.

L'estomac

Les fonctions motrices de l'estomac comprennent, outre une fonction de réservoir, le mélange et le brassage des aliments et la régulation de leur évacuation vers le duodénum. L'essentiel de l'activité motrice de l'estomac a lieu immédiatement après le repas. Les recherches sur la motilité[1] de l'estomac concordent sur la présence "d'ondes électriques lentes" (appelée aussi, dans le contexte de l'EGG, "activité électrique de contrôle" ou "rythme électrique de base"), qui ont une fréquence de 3 cycles par minute et qui sont toujours en phase avec les contractions enregistrées par manométrie (dont l'amplitude moyenne est de l'ordre de 2,5 kPa) [3, 4, 22, 23]. L'amplitude des contractions est déterminée par un deuxième type d'activité électrique, les "potentiels d'action" ("activité de réponse électrique" ou, en anglais, *spikes*). Une étude temps-fréquence publiée par Akin et Sun montre que les potentiels d'action, toujours superposés aux ondes lentes, ont une fréquence de 50-80 cycles par minute [4]. Ces potentiels d'action durent quelques secondes et semblent être plus fréquents en période postprandiale[2]. Chen *et al.* [22] observent aussi des basses fréquences sur les EGG, pendant les contractions. Une explication possible est avancée par les mêmes auteurs dans [23], qui supposent une superposition des ondes

[1] motricité/motilité = n. fém. BIOL. Propriété des êtres vivants et de leurs parties (organes, cellules, organites) de se mouvoir spontanément. © Hachette Livre, 1998

[2] Par période *postprandiale*, ou *digestive*, on entend l'intervalle temporel qui commence immédiatement après un repas et qui est caractérisé par une activité intense de l'estomac. Durant la période *inter-digestive*, qui suit celle postprandiale, l'activité est principalement localisée au niveau de l'intestin grêle.

générées par l'intestin grêle aux ondes lentes de l'estomac. Cette superposition a lieu autant dans les basses fréquences (1 cycle par minute), surtout en période postprandiale, que dans des plus hautes fréquences (12 cycles par minute), entre deux prises alimentaires (quand l'activité de l'estomac est moins importante). Enfin, des études récentes [3] précisent qu'il existe une activité électrique encore plus basse fréquence de l'estomac (1 cycle toutes les 1 à 6 minutes) qui génère des contractions lentes et de faible amplitude (1 à 1,5 kPa).

L'intestin grêle

L'intestin grêle est le segment du tube digestif où l'essentiel des processus de digestion et d'absorption a lieu. La motricité intestinale normale varie fortement par rapport à la prise alimentaire :

- en période postprandiale (digestive), la motricité intestinale présente une forte activité à la fois segmentaire (de mélange) et péristaltique (qui pousse le contenu dans le sens estomac-côlon) ;
- lorsque les phénomènes digestifs sont achevés, l'intestin grêle garde une activité cyclique, appelée Complexe Moteur Migrant (CMM), assurant un flux lent orienté vers le côlon.

Otterson et Sarr, dans [82], décrivent un fonctionnement intestinal qui respecte les mêmes principes que celui de l'estomac : le rythme électrique de base (REB), d'une fréquence de 11 à 13 cycles par minute (voir aussi [23]) dans l'intestin proximal et de 8 à 10 cycles par minute dans l'intestin distal, est surchargé par des potentiels d'action (PA) qui rendent les contractions musculaires efficaces. L'étude citée [82] distingue trois classes principales de contractions intestinales, normales et pathologiques :

1. D'abord, les contractions les plus courantes sont les contractions phasiques individuelles (*individual phasic contractions*) à la fréquence du REB, qui constituent l'activité de base de l'intestin grêle et qui apparaissent à la fois en phase digestive (postprandiale) et inter-digestive. Ce sont probablement ces contractions qui s'organisent parfois en groupes qui durent 1 minute et qui sont séparées par des périodes d'inactivité d'une minute. Ce phénomène est peut-être l'explication de la présence des basses fréquences (1 cycle par minute) dans les enregistrements électrogastrographiques stomacaux de Chen *et al.* [23], cité précédemment. L'activité rythmique décrite ci-dessus est nommée "rythme minute" dans [3].
2. Ensuite, toujours dans le cadre d'une motricité normale, on observe des groupes de contractions organisées, qu'on partage en deux sous-classes : le Complexe Migrant Moteur (CMM), qui caractérise la période inter-digestive, et les contractions migrantes groupées (CMG) (*migrating clustered contractions*), qui se superposent au CMM, mais apparaissent aussi en période postprandiale.

   Le CMM, qui implique parfois aussi l'estomac et se propage jusqu'au début du côlon, est une séquence de contractions qui dure en moyenne entre 90 et 120 minutes [3, 9, 62, 82, 106, 114]. L'intestin grêle est le siège d'une intense activité mécanique caractéristique du CMM, qui affecte à un instant donné tout un segment, puis qui migre sur toute la longueur de l'organe. Lorsque ce front d'activité a atteint la fin de l'iléon, un autre cycle réapparaît. En un segment donné de l'intestin grêle, on enregistre successivement quatre phases :
   - la phase I, de "quiescence", est presque sans contractions et dure en moyenne de 15 minutes [62, 106] à 65 minutes [3, 9], soit de 20% à au plus 50% de la durée totale du CMM ;
   - la phase II, de transit, qui se caractérise par des contractions irrégulières mais de plus en plus fréquentes et qui dure de 30 minutes [9] aux environs de 60 minutes [3, 62, 106], soit de 30% à 50% ;

- la phase III, la plus importante, qui est la phase active du CMM, est caractérisée par des contractions régulières et fortes, à une fréquence de 10-12 par minute, et dure entre 5 et 15 minutes ;
- la phase IV, de courte durée, est une phase de transition entre la phase III et la phase I.

Les CMG durent de 1 à 3 minutes, sont fortes et plus localisées que le CMM (les contractions se propagent sur des distances variant entre 10 et 30 cm). Elles sont plus fréquentes dans la partie distale de l'intestin grêle. Leur rôle principal semble être l'évacuation vers le côlon du contenu du jéjunum [82].

3. Enfin, la troisième classe est constituée des contractions propulsives spéciales, qui apparaissent souvent dans des situations pathologiques. On a aussi deux sous-classes, les contractions géantes rétrogrades, qui favorisent le vomissement, et les contractions géantes migrantes, qui poussent fortement le contenu intestinal vers le côlon. Ce deuxième type de contractions est rencontré parfois dans des conditions normales, en période inter-digestive, mais est plus fréquent en présence d'états pathologiques (diarrhée) ou après l'ingestion de certains médicaments [82].

Le côlon

D'après Thomson et Shaffer, les connaissances sur la motilité du côlon sont moindres que celles sur les autres régions du tractus gastro-intestinal [114], principalement parce que le transit des aliments du cæcum au rectum est un processus lent et irrégulier qui demande quelques jours, il est donc difficilement observable. On sait pourtant que la motricité colique varie fortement avec l'alimentation :

- lors du jeûne, l'activité motrice du côlon est aléatoire et les périodes de silence peuvent être longues. A des intervalles de temps irréguliers (20 à 30 minutes), ces périodes d'inactivité sont interrompues par des bouffées de contractions. Souvent rythmique, cette activité dure quelques minutes et sépare l'organe en segments, sans générer de mouvement du contenu [3, 114] ;
- en période postprandiale, la motricité colique augmente durant 1 à 2 heures [3]. Cette augmentation porte à la fois sur la segmentation et le péristaltisme puisque la défécation est fréquemment déclenchée en période postprandiale. Le terme de "réflexe gastrocolique" est habituellement utilisé pour définir cette stimulation du côlon par la prise alimentaire, même si une activité comparable est observée vers la fin de la nuit, sans aucune stimulation gastrique [3].

Cette description apporte déjà des informations sur les types d'événements que nous risquons de rencontrer. Les variations de fonctionnement de l'appareil digestif se manifestent par des changements des différents indices (énergie, nombre d'événements par minute, etc.), on peut observer une certaine périodicité de l'activité gastro-intestinale et on peut extraire des informations concernant le fonctionnement normal ou pathologique de l'abdomen. Globalement, les connaissances sur l'activité motrice du tractus gastro-intestinal nous permettent de conclure que :

- la motilité des organes abdominaux varie fortement avec le cycle digestif, et que les différences sont nettes entre la période postprandiale (phase digestive), dont la durée varie en fonction du repas et peut dépasser 2 heures, et celle du jeûne (phase inter-digestive), caractérisée par la présence des phénomènes cycliques comme le CMM, installée après la phase digestive et durant jusqu'au repas suivant (voir tableau 1.1) ;
- les mesures externes (non-invasives) de l'activité spécifique d'un organe sont perturbées par l'activité des organes voisins.

	Phase digestive	Phase inter-digestive
Estomac	3 cycles/min	activité faible
Intestin grêle	rythme minute + CMG	CMM
Côlon	réflexe gastrocolique	activité irrégulière

TAB. 1.1 – Motricité du tractus gastro-intestinal en fonction des étapes postprandiales. La phase digestive commence immédiatement après le repas, peut durer plus de 2 heures et est suivie par la phase inter-digestive.

1.1.2 Méthodes d'investigation

Les travaux concernant la motilité abdominale se concentrent, pour la plupart, sur l'aspect médical. Les connaissances dans ce domaine sont peu nombreuses et dépendent fortement de la nature des signaux étudiés : électriques, manométriques, acoustiques, etc. Dans cette section, nous présentons un état de l'art des méthodes d'acquisition et de traitement des signaux traduisant l'activité motrice du tube digestif, en mettant l'accent sur les signaux sonores.

Les signaux abdominaux

Une grande partie des études de la motilité gastro-intestinale proposent d'autres moyens d'investigation que l'auscultation des sons abdominaux. Ces méthodes, dont l'objectif principal est l'exploration fonctionnelle de l'appareil digestif et non pas le diagnostic médical, sont décrites brièvement dans les paragraphes suivants.

L'électrogastrogramme (EGG). Cet examen, similaire à l'électrocardiogramme, est décrit par plusieurs auteurs. Les électrodes sont placées soit directement sur la paroi de l'estomac [4], soit sur la peau du ventre (EGG de surface) [22, 23, 67]. Cette méthode permet de détecter l'activité électrique des muscles lisses de l'appareil digestif. Cependant, elle présente quelques inconvénients importants :

- elle est appropriée, comme son nom l'indique, surtout pour l'activité électrique de l'estomac ;
- dans sa version invasive, elle n'a jamais été appliquée sur des sujets humains ;
- dans le cas de l'EGG de surface, les perturbations dues à l'activité électrique d'autres organes peuvent être importantes et le signal est difficilement interprétable[3].

La manométrie. Une méthode plus ancienne est l'étude des variations de pression à l'intérieur des viscères avec des sondes manométriques [44, 56, 62, 68, 102, 105, 106, 120, 122]. Les contractions des muscles des parois des organes abdominaux sont détectées par des capteurs introduits dans l'estomac et les intestins. Cet examen est sûrement un des plus informatifs, notamment par les indices de motilité qu'il fournit : le nombre, la fréquence et l'énergie des contractions, leurs localisation et propagation, leur évolution par rapport au CMM ou aux différents états physiologiques (après/avant repas, sommeil). Cependant, il présente l'inconvénient important d'être fortement invasif : outre l'inconfort subi par le patient, le fonctionnement même des organes risque d'être perturbé par la présence des sondes.

[3]Néanmoins, différents auteurs ont appliqué cette méthode en intégrant en aval des algorithmes de traitement du signal (par exemple, distribution temps-fréquence de Choi-Williams et réseaux de neurones [22], modèles spectraux et transformation continue en ondelettes [4]).

Les vibrations de surface. A la fin des années 80, Cullen et Campbell ont publié des approches inspirées de l'étude des matériaux [17, 33]. Les vibrations de la surface abdominale (*SVA - surface vibration analysis*[4]) sont enregistrées à l'aide d'un capteur de vibrations piézoélectrique, permettant ainsi de visualiser l'évolution de l'activité gastro-intestinale après un repas standard, ainsi que dans plusieurs cas simulés d'obstruction intestinale ou après ingestion de médicaments. Malheureusement, la bibliographie très succincte ne nous a pas permis d'approfondir ce sujet.

L'échographie. Deux autres études, qui associent l'imagerie médicale aux explorations fonctionnelles, proposent des méthodes ultrasoniques pour l'étude de l'activité gastrique (voir [58] pour une échographie en temps-réel de l'estomac) ou de l'intestin grêle (pour la détection des ondes péristaltiques à l'aide de l'échographie Doppler [51]).

Les sons abdominaux

Une autre méthode d'investigation de l'activité gastro-intestinale est l'auscultation. Elle est courante en clinique et elle est facile à mettre en œuvre. L'auscultation ne perturbe pas les processus physiologiques car elle est complètement non-invasive, ne fait subir aucun inconfort au patient et apporte des informations diagnostiques.

Nous avons choisi de consacrer notre recherche à l'étude des manifestations sonores gastro-intestinales. Cette section est dédiée à la présentation des travaux réalisés dans ce domaine. Après quelques considérations générales sur l'utilité de leur étude, nous présentons les différents protocoles de mesure et méthodes de traitement, analyse et interprétation du signal.

Les premières références bibliographiques datent du début du siècle dernier (Cannon, 1905 [18]). A notre connaissance, la seule étude qui évalue d'une manière quantitative l'utilité diagnostique de l'auscultation des sons abdominaux est en revanche récente (1998) : Gade *et al.* [47] comparent les diagnostics posés par cent médecins à partir d'un ensemble de sons enregistrés et concluent que l'examen est pertinent pour certaines pathologies, mais qu'il est peu fiable pour d'autres. Cependant, l'auscultation abdominale est un examen systématiquement appliqué en clinique [18, 47, 49, 76]. L'enregistrement et le traitement plus ou moins automatique des sons abdominaux ont été proposés par plusieurs auteurs (les méthodes de traitement seront détaillées dans la suite) [7, 15, 30, 31, 32, 39, 41, 48, 84, 98, 99, 110, 116, 117, 119, 121, 122, 126].

Généralement, la bibliographie concorde sur l'appréciation de l'efficacité de cet examen : dans l'état actuel, il est difficilement exploitable du point de vue médical, surtout à cause des difficultés d'acquisition et d'interprétation [15, 116]. De plus, les opinions des auteurs varient : selon certains, les caractéristiques des sons abdominaux et même leur absence ne permettent pas d'établir le diagnostic d'un état particulier, sauf, d'après Thomson et Shaffer, les bruits très aigus de l'occlusion aiguë de l'intestin grêle [114]. En fait, comme plusieurs auteurs le soulignent, il est probablement plus utile de déterminer la façon dont les sons évoluent sur une longue durée d'auscultation [49, 114, 127], concordante avec la physiologie gastro-intestinale.

Acquisition. Protocoles de mesure. Pour l'acquisition des sons abdominaux, le capteur courant en clinique est le stéthoscope acoustique ou, dernièrement, le stéthoscope électronique. Ces capteurs sont aussi utilisés pour l'enregistrement automatique des sons : Garner et Ehrenreich [48] utilisent une tête de stéthoscope acoustique à laquelle ils attachent un microphone, alors que, récemment, Craine *et al.* [30, 31, 32] utilisent des stéthoscopes électroniques commer-

[4]Une technique analogue est utilisé par Royston et Mansy [75, 95] pour étudier la propagation des sons dans l'organisme humain.

ciaux[5]. Souvent, dans le cas de l'enregistrement automatique, on rencontre différents types de microphones directement placés sur la surface abdominale.

En ce qui concerne la position des capteurs, on retrouve des localisations communes dans la quasi-totalité des publications : en région épigastrique (au niveau de l'ensemble pylore-duodénum) pour les contractions stomacales, dans le cadran inférieur gauche pour l'intestin grêle, et dans le cadran inférieur droit (au niveau de la valve iléo-cæcale) pour la fin du jéjunum et le côlon ascendant. L'auscultation simultanée en plusieurs endroits proposée par Craine [32] respecte cette distribution des microphones. Une autre étude multi-voies, réalisée par Garner et Ehrenreich [48]. utilise six capteurs, dont trois placés aux endroits énumérés ci-dessus. Dans cette étude, deux capteurs supplémentaires sont rajoutés au niveau des angles coliques (les angles droits faits par le côlon au début et à la fin de sa partie transverse) et un troisième est dédié au bruit ambiant.

Les méthodes de conditionnement du signal varient beaucoup. En fait, comme une grande partie des études date de plus de 20 ans, le traitement est souvent analogique et la chaîne d'instrumentation n'est composée que de filtres passe-bas ou passe-bande, utilisés pour éliminer les sons cardiaques et pulmonaires de plus basse fréquence. Les études plus récentes recommandent de numériser le signal, avec des fréquences d'échantillonnage qui passent de 200 Hz. [110] à 1000 Hz. [116], 4000 Hz. [15], 4096 Hz. [74], 5000 Hz. [34, 54] et 8000 [30] ou 11025 Hz. [31, 32]. La plupart des auteurs considèrent que la fréquence maximale des sons abdominaux est en-dessous de 1000-1500 Hz. [15, 32, 74, 84, 116, 126], même si d'autres valeurs sont mentionnées dans [34] (2500 Hz.) et [48] (5000 Hz.).

Enfin, pour enregistrer les bruits abdominaux, chaque auteur propose son propre protocole de mesure. Les durées d'enregistrement varient de quelques minutes [30, 31, 49] à plusieurs heures, distribuées sur une journée complète [7]. La standardisation des repas est courante [7, 30, 31, 32, 116, 119] mais pas obligatoire [15, 121]. Selon certaines sources bibliographiques (Thomson et Shaffer [114]), la palpation ou la percussion abdominales influencent la génération des bruits abdominaux, qui devraient donc être recherchés avant ces manœuvres cliniques. Cette affirmation est contestée par d'autres études (Sandler *et al.* [97], Angtuaco *et al.* [5]).

Traitement des signaux et des événements. Les méthodes automatiques de traitement du signal pour les sons abdominaux sont assez peu développées : dans son article de synthèse sur les méthodes de traitement des signaux de motricité digestive qui date de 1988, Rozé [96] ne mentionne aucun système utilisant les sons. En effet, probablement à cause de la difficulté du traitement des signaux sonores mais aussi à cause de l'ambiguïté des informations contenues dans les sons abdominaux, pratiquement aucun appareil de ce type n'est utilisé couramment en routine clinique ou en recherche médicale.

Le schéma classique de traitement de ce type de signal comporte plusieurs étapes :
- pré-traitement (débruitage/segmentation) ;
- analyse du signal et des événements (définition et extraction des caractéristiques informatives, classification) ;
- interprétation des résultats et application.

Dans la bibliographie étudiée, le **pré-traitement** consiste surtout en la segmentation des signaux en vue de l'extraction des événements informatifs et il inclut rarement le débruitage, à l'exception notable des travaux de Hadjileontiadis *et al.* [54]. Les méthodes de segmentation varient, surtout en fonction de leurs dates de parution : les premières méthodes sont analogiques

[5]En général, les caractéristiques acoustiques d'un stéthoscope électronique sont déterminées plus par l'électronique intégrée que par la forme et les matériaux utilisés pour sa construction, comme dans le cas des stéthoscopes acoustiques classiques.

et font appel souvent à l'expertise visuelle et auditive d'un médecin [7, 18, 39, 41, 48, 121]. D'autres méthodes proposent un seuillage d'amplitude, analogique ou numérique, sur le signal [84, 110, 119] ou sur des versions filtrées du signal [30, 31, 32]. Enfin, d'autres algorithmes proposent un seuillage sur l'enveloppe du signal, analogique ou calculée à l'aide d'une fenêtre glissante ou par transformée de Hilbert [34, 74]. Évidemment, le type de pré-traitement choisi dépend aussi des objectifs et des méthodes mises en œuvre dans les étapes suivantes.

L'analyse des événements segmentés passe par une extraction de caractéristiques, telles que la durée, l'énergie, la fréquence principale (*pitch*). La durée et l'énergie sont facilement obtenues à partir du signal segmenté. Le contenu fréquentiel est calculé soit analogiquement, par bancs de filtres [121], soit par transformée Fourier discrète [15, 32, 34, 74, 116].

Il est possible d'associer à ces caractéristiques des informations plus détaillées sur la structure d'un événement : d'une manière simple, les auteurs de [48] rappellent qu'un événement peut être constitué d'un son abdominal isolé (*click*), mais il peut aussi contenir une succession rapide de sons individuels (*multi-click*).

Enfin, l'approche multi-voies introduite par Garner et Ehrenreich [48] est utilisée pour étudier la propagation et la localisation des sons le long des intestins. Une autre méthode multi-voies, fondée sur une triangulation, est utilisée par Craine *et al.* [32] pour réaliser une carte de l'activité abdominale, normale et pathologique. L'information obtenue par localisation est ensuite associée aux autres caractéristiques pour analyser le fonctionnement du tractus gastro-intestinal.

La description des sons abdominaux par leurs caractéristiques est utilisée pour leur classification. Cette classification est souvent subjective, avec des frontières entre les classes définies par des seuils utilisateur (sons longs/courts, sons de fréquence basse/haute, etc.).

L'analyse globale des signaux est une autre voie, exploitée dans plusieurs études. Cette approche peut se faire sans aucun pré-traitement : Yoshino *et al.* [126] proposent un découpage des signaux en séquences de 30 secondes, suivi d'une transformation de Fourier appliquée sur chacune de ces parties. Un traitement global permet ensuite d'étudier l'évolution des caractéristiques fréquentielles des signaux enregistrés.

Néanmoins, tous les autres auteurs utilisent la segmentation, car un des indicateurs principaux de la physiologie abdominale est le nombre de sons par unité de temps. Ce n'est pas le seul indice de l'activité gastro-intestinale :

- certains auteurs [7, 116] définissent des indices de motilité à partir de l'amplitude moyenne des sons par unité de temps (l'intégration de l'énergie du signal sur des intervalles de temps a été introduite, en analogique, en 1955 par Farrar et Ingelfinger [41]) ;
- la durée des intervalles de silence entre deux sons est l'indicateur principal utilisé par Craine *et al.* [30, 31, 32], qui montrent une corrélation avec le nombre de sons par intervalle de temps ;
- le rapport des durées totales des sons et des intervalles de silence est aussi proposé par Vasseur [110, 119].

Enfin, plusieurs études combinent **l'analyse des événements** et **l'analyse globale des signaux** : les indices d'activité énumérés dans le paragraphe précédent sont appliqués non pas au signal en entier, mais aux classes de sons abdominaux. Ainsi, Bray [15] calcule le nombre d'événements en fonction de leurs fréquences, Tomomasa *et al.* [116] proposent une représentation 3D du nombre de sons rapporté à leur durée et à leur fréquence principale, Craine *et al.* [32] calculent un spectre moyen des événements, Vasseur *et al.* [119] une énergie moyenne etc.

Application médicale et interprétation. Les premières applications médicales de l'auscultation abdominale sont en fait les premières observations sur le fonctionnement rythmique des

organes de l'appareil digestif (Cannon, 1905 [18]), ce qui revient à étudier les variations des indices d'activité définis antérieurement. On remarque déjà la fréquence de base de l'estomac (3 cycles par minute), ainsi qu'une approximation de la rythmicité des mouvements de segmentation de l'intestin grêle (7 à 12 contractions par minute). La différence entre la phase digestive et la phase inter-digestive n'est pas encore clairement affirmée, mais l'hypothèse de l'existence du CMM est déjà avancée. Ce comportement rythmique d'une partie du tractus digestif (estomac et intestin grêle) et ses manifestations sonores sont confirmés aussi par de nombreux auteurs [41, 48, 49, 76], même s'ils relativisent la dépendance stricte entre les contractions et les manifestations sonores. Enfin, dans une étude récente, Tomomasa *et al.* [116] montrent une liaison entre les phases du CMM et l'activité sonore.

Cependant, les publications traitant des sons abdominaux ne sont pas uniquement dédiées à l'étude de la physiologie normale de l'appareil digestif. Différents auteurs proposent une description de la motilité gastro-intestinale (normale, hyper- ou hypo-activité), quantifiée parfois par des indices de motilité (indices d'activité) [7, 116]. Cette description est ensuite reliée à un état physiologique : diverses pathologies [2, 7, 30, 31, 39, 49, 110], phases du cycle digestif [41, 116, 119], traitements pharmaceutiques ou chirurgicaux [41, 48, 122]. Dans plusieurs développements, outre la caractérisation globale des signaux sonores abdominaux, une analyse individuelle des sons est prise en compte :

- les caractéristiques fréquentielles et/ou temporelles des sons sont associées à un état précis : l'occlusion intestinale, par exemple, est signalée soit directement par des bruits aigus, métalliques [2, 39, 41, 49], soit par un spectre particulier d'une partie du signal enregistré [126] ;
- les phases du cycle digestif (avant/après repas) semblent aussi influencer la nature des sons abdominaux : leur distribution fréquentielle est différente [15], ainsi que leur durée moyenne et leur énergie [119] ;
- une localisation des sons en fonction de leurs caractéristiques est envisageable : les sons longs sont souvent associés au péristaltisme et aux contractions stomacales, alors que les sons courts sont associés plutôt aux mouvements de segmentation ou aux contractions du côlon [15, 18, 39, 121]. Pourtant, d'après Milton, la localisation des sons par leurs caractéristiques reste un sujet controversé [76].

Certains auteurs [18, 49, 76, 84] proposent des modèles physiques de génération des bruits abdominaux, en insistant généralement sur la nécessité de la présence de gaz (ce qui est contesté par [121]), mélangé dans des proportions variables aux liquides. D'autres études récentes en acoustique [75, 95] introduisent des modèles de propagation des sons dans les poumons et l'abdomen : le diagnostic du pneumothorax[6] ou du pneumopéritoine[7] est fait à l'aide d'une modélisation de l'organisme comme un milieu visco-élastique présentant des anisotropies (poches d'air) détectables par l'étude de la propagation des sons. Cette modélisation acoustique dépasse le cadre de notre travail.

1.1.3 Conclusion

Cette présentation des aspects physiologiques et médicaux permet de supposer l'existence d'un lien entre le fonctionnement de l'appareil digestif et ses manifestations sonores. Même si la communauté médicale semble partagée sur l'utilité et le pouvoir diagnostic d'une auscultation

[6] Introduction spontanée ou provoquée d'air entre les deux feuillets de la plèvre. ©Hachette Livre, 1998

[7] Présence de gaz entre les feuilles du péritoine ; elle peut être spontanée, lors de la perforation d'un organe creux (appendicite, ulcère gastro-duodénal), ou artificielle, provoquée en vue d'un examen radiologique des organes abdominaux. ©Hachette Livre, 1998

automatisée, nous pouvons espérer que l'étude des sons abdominaux enrichira les connaissances sur la physiologie gastro-intestinale normale et pathologique. Cet objectif nous conduit vers la problématique scientifique qui lui est associée. Pour pouvoir déduire cette problématique, qui sera détaillée dans la section suivante (1.2), nous présentons ici les conclusions principales de la partie biomédicale.

Premièrement, le fonctionnement **normal** du tractus gastro-intestinal est rythmique, du moins pour sa partie proximale (estomac et intestin grêle). Ce rythme varie en fonction de la phase du cycle digestif (phase digestive – immédiatement après le repas, et phase inter-digestive). Cette activité rythmique se manifeste toujours électriquement et par variations de pression et, souvent, par génération de sons abdominaux. Ces sons, qui peuvent éventuellement être séparés en quelques classes selon leurs caractéristiques, leur origine et/ou leur mécanisme de génération, sont considérés comme normaux du point de vue médical. Dans la littérature, la classification des sons abdominaux est souvent subjective et fait appel à des classes peu précises : les gargouillements, ou les borborygmes, peuvent être des sons explosifs, des clapotis, des sons métalliques, des *clicks*, des *stacatto pops*, etc. [18, 39, 41, 49, 126].

Les différences de motilité entre les phases du cycle digestif, détectables par des moyens d'investigation comme l'électrogastrogramme ou la manométrie, semblent mesurables aussi par l'intermédiaire des sons, sans pour autant avoir des sons spécifiques : la différence se fait sur la globalité du signal, en utilisant des indices d'activité comme l'énergie et/ou le spectre du signal, ou encore le nombre de sons par minute, le nombre de sons par classe (les critères de séparation en classes restent à définir, mais ils comprennent des paramètres comme la durée ou la fréquence principale du son). Il faut remarquer que cette approche impose la nécessité d'un pré-traitement, car elle pose un problème de segmentation, c'est-à-dire de définition d'un son abdominal[8].

Deuxièmement, un état **anormal** du tractus digestif (pathologique et/ou pendant ou après une intervention médicale ou chirurgicale) se manifeste soit par une motilité différente, soit par des sons particuliers. Les désordres de motilité peuvent être mesurés et détectés par les mêmes méthodes que celles utilisées pour la différenciation entre les phases du cycle digestif, décrites précédemment. En revanche, les sons particuliers (spécifiques d'une occlusion, par exemple), sont classés différemment, en fonction de leurs caractéristiques (fréquence, durée, puissance, localisation).

Enfin, il faut rappeler que la plupart des sources bibliographiques soulignent la nécessité d'une auscultation abdominale respectant trois conditions essentielles :

- l'écoute doit être faite sur de longues durées, autant pour apprécier la motilité et le péristaltisme, que pour capturer les événements particuliers d'un phénomène pathologique ;
- il est préférable d'enregistrer l'activité abdominale dans des conditions standardisées (même repas et mêmes horaires) ;
- la localisation des capteurs est importante et doit être faite en tenant compte à la fois des objectifs et des zones du tractus digestif les plus actives, du point de vue sonore.

Les possibles objectifs biomédicaux qui découlent de cette présentation sont divers :

- la validation de l'auscultation abdominale comme méthode d'investigation fonctionnelle, en la comparant aux connaissances acquises par d'autres méthodes, ce qui permettra une utilisation courante fiable, non-invasive et suffisamment informative pour orienter les démarches médicales ultérieures ;
- la caractérisation des sons abdominaux et leur classification, suivies d'une étude sur la corrélation entre les classes et les modes de fonctionnement (phases du cycle digestif, états pathologiques, interventions thérapeutiques) ;

[8] Un son abdominal sera nommé aussi, au cours de cette thèse, un événement.

- une localisation des zones actives du point de vue sonore, sous l'hypothèse que l'activité sonore témoigne de l'activité motrice, et la création d'une carte fonctionnelle abdominale ;
- une association des deux derniers points, pour étudier une éventuelle corrélation entre les types de sons et leur localisation (et donc l'organe ou le mécanisme physiologique qui génère les sons).

1.2 Problématique scientifique

La première section de ce chapitre (1.1), après une présentation de l'anatomie et de la physiologie du tractus digestif, montre l'intérêt de l'étude des bruits abdominaux pour les différentes applications cliniques et de recherche biomédicale. Comme nous l'avons indiqué dans la section précédente (1.1.2), cette étude comporte plusieurs étapes (voir figure 1.2) :

- l'acquisition multi-voies des signaux sonores abdominaux (non-stationnaires, bruités et de longues durées), ce qui pose le problème d'une instrumentation adaptée ;
- un pré-traitement adapté à son tour à la nature des signaux enregistrés, ce qui implique des phases successives de débruitage, détection, segmentation et localisation spatiale des événements ;
- un traitement spécifique, comportant deux volets :
  - une analyse globale, portant sur la motilité digestive, qui vise à décrire et à classer les signaux à l'aide de certains indices d'activité et de leur évolution temporelle ;
  - une analyse individuelle, événement par événement, visant l'extraction des caractéristiques physiques (durée, fréquence, énergie), suivie par la classification ;
- l'utilisation de l'information extraite lors des étapes précédentes pour l'analyse des modes de fonctionnement de l'abdomen et l'aide au diagnostic.

1.2.1 Signal et instrumentation

Les signaux d'intérêt, appelés *phonoentérogrammes*, sont des signaux sonores de basse fréquence (d'après les auteurs cités précédemment, une bande de fréquence entre 80 et 2500 Hz. est largement suffisante), hautement non-stationnaires, dont l'information est répartie sur de longues durées. A ce signal se superpose, dans la même bande de fréquence, un bruit dû à la fois à l'instrumentation et à d'autres sources sonores internes (cœur, poumons) ou externes. De plus, comme les enregistrements sont faits à l'aide de plusieurs capteurs, un même son abdominal peut se propager sur plusieurs voies. Le choix de l'instrumentation, ainsi que les méthodes de traitement ultérieur, sont déterminés par ces caractéristiques.

Une partie des problèmes soulevés par l'enregistrement des sons abdominaux est commune à toute acquisition de signaux sonores : le choix et le positionnement du ou des capteurs, le choix des caractéristiques de la chaîne de traitement, le débruitage, la reproductibilité des mesures, le stockage des données. Cependant, comme les signaux d'intérêt dans le cadre de ce travail sont les sons physiologiques, une précision nous semble nécessaire : la quasi-totalité des informations que l'on a sur ces signaux est obtenue à travers le stéthoscope, qui est l'instrument de base de l'apprentissage médical et de la routine clinique. Par conséquent, il est le seul moyen de mesure qui permette la validation médicale indispensable pour notre application. De plus, par sa construction, le stéthoscope assure une forte diminution des perturbations extérieures dues au bruit ambiant. Dans un premier temps, notre capteur est constitué d'une tête de stéthoscope équipée d'un microphone.

Si ce choix assure une interprétation et une validation médicales aisées, il réclame une étude du capteur, qui joue le rôle d'un filtre-amplificateur mécanique dont la fonction de transfert doit

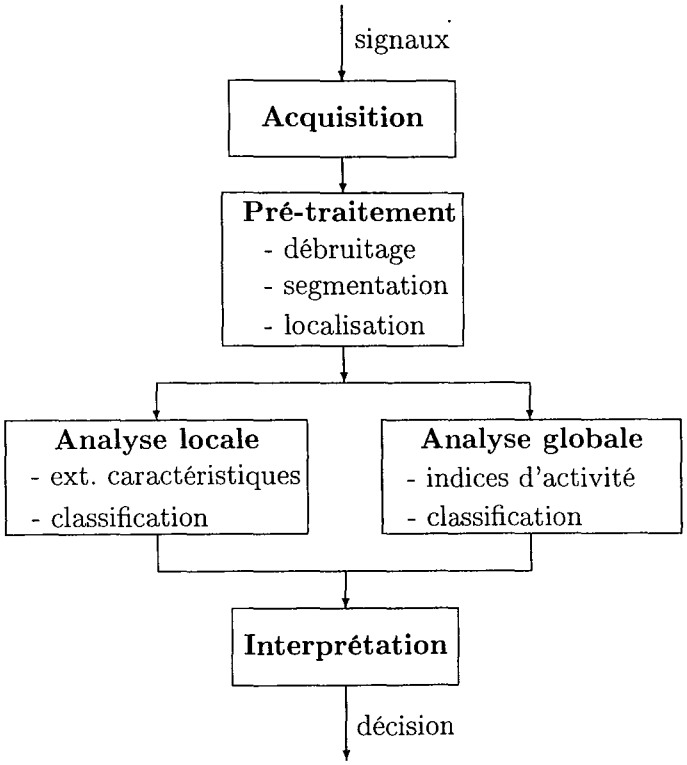


FIG. 1.2 – Schéma général des étapes de traitement constituant un outil d'aide à la décision.

être identifiée.

Un deuxième problème vient de la fixation des têtes de stéthoscope. En fait, la pression du stéthoscope sur la peau risque d'influencer ses propres caractéristiques de filtrage et donc doit être à la fois relativement contrôlée et prise en compte dans tout traitement ultérieur, pour assurer une reproductibilité des mesures et des résultats.

Le positionnement des têtes de stéthoscope est également important à établir, car les sons se propagent dans l'abdomen et leurs caractéristiques sont modifiées par le milieu et la distance de propagation.

Un dernier problème lié à l'instrumentation est la détermination de la fréquence d'échantillonnage du signal, qui doit prendre en compte à la fois le contenu fréquentiel du signal et les capacités de stockage et de traitement des signaux de longue durée.

Enfin, pour valider notre travail, un protocole d'acquisition clair et précis doit être mis en place. Le respect d'une certaine rigueur dans les horaires d'alimentation et d'enregistrement, ainsi que dans le placement des stéthoscopes, est indispensable pour une première évaluation de nos recherches.

1.2.2 Pré-traitement

Le pré-traitement des signaux sonores abdominaux comporte plusieurs étapes : le débruitage, la détection et la segmentation des événements, suivis par leur localisation spatiale.

L'élimination du bruit et la segmentation sont liées et dépendent de la définition que l'on donne au bruit et au signal informatif à extraire. Pour séparer le bruit du signal, on doit avoir un critère de séparation, qui implique des connaissances plus ou moins approfondies sur ces deux composantes du signal mesuré.

La définition la plus simple du bruit est négative : on appelle bruit toute composante du signal mesuré qui n'est pas informative pour l'application. Pour être plus précis, on peut néanmoins séparer deux grandes classes de bruit : le bruit de mesure, dû essentiellement à l'instrumentation, et les artéfacts (les signaux parasites), générés par d'autres sources sonores que celles d'intérêt.

Dans le cadre de notre application, la plupart des artéfacts sont des bruits ambiants et des sons physiologiques générés par d'autres organes, comme le cœur et les poumons. Ces artéfacts peuvent être éliminés ou diminués par un pré-traitement adéquat : une première approche ne nécessite pas de connaissance *a priori* sur le contenu fréquentiel, mais considère le signal informatif et le bruit comme des processus stochastiques. Il s'agit des techniques de décorrélation de type filtrage de Wiener ou filtrage adaptatif (pour prendre en compte les signaux non-stationnaires) qui demandent l'accès à la source du bruit [78, 125]. Dans le cadre de notre application, cette méthode a été proposée par Suzuki *et al.* [111] pour les auscultations pulmonaires et par Mansy et Sandler [72, 73], pour les enregistrements abdominaux. Dans une étude antérieure (DEA [85]), nous avons testé et adapté cette technique aux signaux acquis avec notre chaîne d'instrumentation.

Nous n'allons pas traiter ici ce sujet : même s'il est indispensable dans l'application clinique, notre préoccupation première est la validation de l'utilité de la phonoentérographie, l'application en routine médicale étant pour l'instant reportée à des travaux futurs. De plus, le choix du capteur nous permet d'atténuer fortement les perturbations extérieures. Ainsi, comme les enregistrements ont été effectués dans un environnement relativement silencieux, nous n'avons utilisé ni le filtrage adaptatif ni d'autres techniques pour éliminer le bruit ambiant, considéré inexistant. Nous avons également travaillé sur des enregistrements non perturbés par le bruit cardiaque : l'apparition de ce type d'artéfact est liée à la taille de la personne et affecte plus souvent les phonoentérogrammes des enfants. Quant aux autres artéfacts possibles (frottements des stéthoscopes contre la peau, par exemple), ils seront traités dans une section qui leur est dédiée (4.2).

Néanmoins, il est nécessaire d'extraire les événements informatifs du signal enregistré et d'éliminer le bruit de mesure. Ce bruit est essentiellement dû aux effets de cavité acoustique de la tête de stéthoscope, mais aussi aux bruits basse fréquence dus aux cavités internes de l'organisme humain, au flux sanguin, à la respiration etc. Ces bruits se trouvent dans la même bande fréquentielle que le signal d'intérêt. De ce fait, le filtrage fréquentiel classique, linéaire, qui suppose une différence entre les contenus fréquentiels du signal informatif et du bruit, est inapplicable.

Si les approches décrites précédemment partent d'hypothèses sur la nature du bruit, utilisées pour le séparer du signal informatif, d'autres méthodes font, au contraire, des hypothèses sur le signal informatif. Il s'agit, notamment, de l'approche proposée par Coifman et Wickerhauser [28, 29] et Donoho et Johnstone [37, 38] dans leurs travaux sur le débruitage par ondelettes. Pour appliquer ces méthodes (qui peuvent être considérées comme des filtrages non-linéaires), les signaux sont représentés dans un autre domaine (temps-échelle) par une transformation en ondelettes. Les auteurs supposent que le signal informatif est corrélé avec une forme d'onde prédéfinie (l'ondelette), qui est suffisamment "lisse", à la différence du bruit supposé aléatoire et décorrélé.

Notre travail se place dans ce même cadre, qui offre l'avantage de prendre facilement en compte la nature non-stationnaire du signal qui nous intéresse. De plus, la décomposition en ondelettes nous fournit des informations utiles pour les étapes ultérieures de classification et d'aide au diagnostic, à la fois sur la structure temporelle et le contenu fréquentiel des événements.

Un avantage important des méthodes de débruitage par ondelettes est de permettre directement la détection des événements et donc une segmentation initiale du signal. Cette segmentation peut être améliorée en incluant d'autres critères, énergétiques, fréquentiels et/ou de localisation, qui peuvent à leur tour être extraits de la décomposition en ondelettes.

Le choix de la représentation du signal dans le domaine temps-échelle (ondelettes) soulève plusieurs questions, qui doivent être traitées conjointement :

- quelles sont la méthode et la base de décomposition en ondelettes les mieux adaptées ;
- comment construire ou améliorer une méthode de débruitage et de segmentation pour les signaux d'intérêt ;
- comment extraire les caractéristiques des événements segmentés et comment les intégrer dans la dernière étape de pré-traitement, la localisation spatiale ?

Le choix de la méthode de décomposition (continue ou discrète) et de la base (orthogonale ou pas) influencent les étapes suivantes, du débruitage à l'extraction des caractéristiques. La non-redondance et l'orthogonalité de la décomposition permettent l'application d'algorithmes performants, tant pour la transformation en ondelettes (l'algorithme "à trous" de Mallat [71]), que pour la détection et le débruitage (l'interprétation énergétique propre aux décompositions orthogonales facilite l'application de certains algorithmes de débruitage qui font référence, par exemple, à l'énergie du bruit). La conservation de l'énergie spécifique aux décompositions orthogonales offre un autre avantage intéressant : l'énergie des événements segmentés peut être estimée localement en temps et en fréquence grâce aux coefficients de la décomposition en ondelettes. Nous avons choisi de travailler dans une base orthonormée : sous cette hypothèse, nous testons plusieurs bases et plusieurs profondeurs de décomposition, afin de trouver la décomposition la plus adaptée.

Les méthodes de détection/segmentation et débruitage par ondelettes présentées dans la littérature sont, dans leur quasi-totalité, développées pour les bases d'ondelettes orthogonales. Néanmoins, elles n'ont pas été appliquées, à notre connaissance, dans le domaine de la phonoentérographie, à l'exception de l'algorithme proposé par Hadjileontiadis *et al.* [54], qui est le point de départ de nos travaux sur le débruitage. Ces méthodes doivent être testées, adaptées et éventuellement améliorées pour notre application.

Après segmentation, chaque événement est caractérisé par sa propre décomposition en ondelettes qui offre une information conjointe en temps et en fréquence. En la comparant aux connaissances *a priori* sur les sons abdominaux, cette information est utilisable pour la détection des artéfacts et la re-segmentation des signaux.

La localisation spatiale apporte une information supplémentaire sur chaque événement et peut contribuer à la construction d'une carte de l'activité abdominale. Afin de construire cette carte, plusieurs modèles de propagation intra-abdominale des sons peuvent être proposés (milieu isotrope [32], propagation sous contraintes anatomiques [48]). La pertinence de ces modèles doit être vérifiée, afin de proposer la méthode la plus adaptée à la localisation des sources sonores intra-abdominales.

1.2.3 Analyse des signaux et des événements

En suivant la conclusion de la section précédente (1.1), le traitement des données en vue d'une interprétation physiologique et médicale se divise en deux grandes parties : l'analyse globale, axée sur des séquences de différentes durées extraites du signal enregistré, et l'analyse locale, concernant les événements segmentés.

La première approche, le **traitement global**, nécessite d'abord la définition des signatures qui puissent caractériser les séquences de phonoentérogramme. Il s'agit de plusieurs indices d'activité calculés par unité de temps : nombre d'événements, rapport silence/sons abdominaux, intervalle de temps moyen entre deux événements, énergie du signal, fréquence moyenne, etc. A l'aide de ces indices, on cherche des perturbations ou des modifications de motilité gastro-intestinale. Les problèmes principaux soulevés par cette approche sont :

- le choix des indices d'activité les plus pertinents pour la description de la physiologie digestive, ou éventuellement d'une combinaison d'indices ;
- l'interprétation des variations des indices d'activité et la validation de cette interprétation.

La deuxième approche, le **traitement local**, implique le choix et l'extraction des caractéristiques significatives des sons abdominaux, comme la durée, l'énergie (la puissance, l'intensité acoustique), la fréquence principale, etc.

Les indices d'activité et les caractéristiques des événements sont des variables nécessaires pour décrire respectivement les séquences de phonoentérogramme et les sons abdominaux individuels. L'analyse individuelle monodimensionnelle de ces variables, le suivi de leur évolution temporelle, voire les différences en fonction de la localisation, peuvent nous fournir un premier lot d'informations sur le fonctionnement de l'appareil digestif.

De plus, les descriptions à l'aide des variables explicatives devraient aboutir à des classifications multidimensionnelles et à des interprétations médicales.

La classification, à la fois des séquences de phonoentérogramme et des sons individuels, est un problème difficile, surtout à cause de l'absence ou de la subjectivité des connaissances *a priori*. Dans ce contexte, l'utilisation d'un algorithme de classification non-supervisée (la recherche des classes "naturelles") semble appropriée. Le choix de cet algorithme est un des points difficiles de cette recherche.

Grâce à un protocole d'enregistrement dédié (enregistrements multi-voies de longue durée), nous pouvons espérer que la répartition en classes obtenue par la classification non-supervisée soit représentative de l'ensemble des signaux sonores abdominaux. Dans ce cas, elle peut être utilisée comme un étalon, comme un ensemble d'apprentissage pour une deuxième phase de classification, cette fois supervisée. L'objectif de cette étape est double :

- la classification supervisée de séquences temporelles successives extraites des phonoentérogrammes devrait fournir une information sur l'évolution de l'activité gastro-intestinale au cours du temps. Par exemple, les sons ou les séquences de phonoentérogramme enregistrés en phase postprandiale devraient appartenir à des classes différentes de ceux obtenus en phase inter-digestive ;
- la classification supervisée vise à classer les nouveaux signaux n'appartenant pas à la base de données utilisée pour la création de l'ensemble d'apprentissage, afin de détecter des éventuels états pathologiques.

Les étapes d'analyse des phonoentérogrammes et des événements individuels sont fortement reliées. En fait, toute extraction de caractéristiques est validée par la cohérence des résultats de la classification qui suit et qui les utilise. L'interprétation médicale des résultats de cette classification, qui constitue la validation finale du projet, devrait permettre d'analyser plus profondément le fonctionnement de l'appareil digestif, voire de certaines de ses parties, par le biais de la localisation.

1.3 Conclusion

Ce chapitre présente les aspects physiologiques et médicaux concernant le fonctionnement de l'appareil digestif et ses manifestations sonores, ainsi que les méthodes d'investigation existantes. Parmi les méthodes d'investigation fonctionnelle, l'auscultation des sons abdominaux est la plus ancienne et la plus utilisée en routine clinique. Même si la communauté médicale semble partagée sur l'utilité d'une auscultation automatisée (peut-être parce qu'elle n'a pas l'expérience d'une auscultation de longue durée), l'étude bibliographique précédente permet de supposer l'existence d'un lien entre la physiologie de l'abdomen, normale ou pathologique, et ses manifestations sonores. L'étude de ce lien est l'objectif médical fondamental de nos recherches.

L'objectif applicatif de cette thèse est la création d'une "boîte à outils" constituée de méthodes adaptées au traitement et à l'analyse des phonoentérogrammes. Nous nous proposons de valider les méthodes mises en œuvre en tant que moyens d'investigation de la physiologie normale de l'appareil digestif, sachant que les mêmes algorithmes devraient être applicables à l'étude des signaux pathologiques et au diagnostic médical.

La dernière partie de ce chapitre est consacrée à la problématique scientifique, associée à la fois aux phénomènes biologiques gastro-intestinaux et au traitement des signaux générés par ces phénomènes. Les étapes de traitement définies sont l'acquisition, le pré-traitement et l'analyse des signaux recueillis.

L'acquisition. L'étude des caractéristiques du signal sonore abdominal nous conduit à définir une instrumentation adaptée, qui doit assurer des acquisitions multi-voies, de longue durée, robustes et qui doit fournir des signaux facilement interprétables par un médecin.

Le pré-traitement. Le débruitage, suivi par la détection, la segmentation et la localisation spatiale des sons abdominaux a pour but d'extraire l'information contenue dans les phonoentérogrammes.

L'analyse des signaux. Les deux volets de cette dernière étape sont :
- la caractérisation globale et la classification des séquences des phonoentérogrammes à l'aide des indices d'activité ;
- la caractérisation locale et la classification des sons abdominaux segmentés à partir de l'information fréquentielle et temporelle contenue.

La suite du manuscrit va s'attacher à décrire tous les points listés ci-dessus. Une attention particulière sera toutefois portée à l'étape de pré-traitement, et en particulier à l'algorithme de détection et débruitage par ondelettes.



Acquisition des sons : signal et instrumentation

La première étape à franchir dans le développement d'un nouvel outil d'investigation clinique est celle de l'acquisition des signaux. Si notre travail principal concerne le développement des méthodes de traitement et d'analyse numériques du signal, il est nécessaire de s'assurer de la validité des signaux à numériser. L'exigence d'obtenir des signaux de bonne qualité, reflétant le plus exactement les sons générés, ainsi que la reproductibilité des mesures, sont indispensables à tout développement ultérieur. Même si notre recherche n'est pas axée sur l'instrumentation, nous avons considéré nécessaire une étude dans ce domaine. Évidemment, pour établir un protocole de mesure et pour choisir les éléments de la chaîne d'instrumentation, il faut caractériser d'abord les signaux d'intérêt. Ce chapitre commence par la description des signaux sonores abdominaux, tels qu'ils sont présentés dans la bibliographie et selon nos constats. La caractérisation des signaux est suivie par la présentation des protocoles de mesure utilisés dans la littérature et par la description du protocole que nous proposons. Enfin, nous décrivons et choisissons les éléments de notre chaîne d'instrumentation, nous proposons une méthode adaptée à sa caractérisation et nous présentons les résultats qui justifient ces choix.

2.1 Signal

Dans cette section, nous décrivons les signaux sonores abdominaux, appelés aussi *phonoentérogrammes*, tels qu'ils sont présentés dans la littérature. Ensuite, nous confrontons cette description à nos constats expérimentaux.

Dans le cadre de l'analyse des sons abdominaux, les signaux d'intérêt sont caractérisés par une succession d'événements isolés et courts, clairement séparés, dont le contenu fréquentiel est relativement pauvre. Ces signaux peuvent être décrits de deux manières :

1. Concernant l'aspect temporel de longue durée, les premiers constats portent sur la rythmicité des sons (périodes actives cycliques de 3 à 12 par minute, en fonction du lieu de leur génération) [18, 41, 49].

   Plusieurs études définissent différents indices d'activité afin de fournir des évaluations quantitatives plus précises sur les phonoentérogrammes. Un des plus courants indices est le nombre de sons par minute. Cet indice d'activité varie énormément en fonction de la définition d'un son et/ou de la méthode utilisée : Arnbjörnsson [7] parle de moyennes de 6

à 25 sons/minute au cours du cycle digestif, Craine *et al.* [32] de 6 à 117 en fonction de la pathologie et de la localisation, alors que Hadjileontiadis *et al.* [54] trouvent de 120 à 1200 sons/minute en fonction aussi de la pathologie. Un autre indice utilisé est la durée totale des sons, ou encore le rapport entre la durée des sons et celle des périodes de silence [110, 119], qui montre que les signaux sonores abdominaux sont creux, avec une durée totale des sons informatifs généralement inférieure à 2% de la durée du signal enregistré.

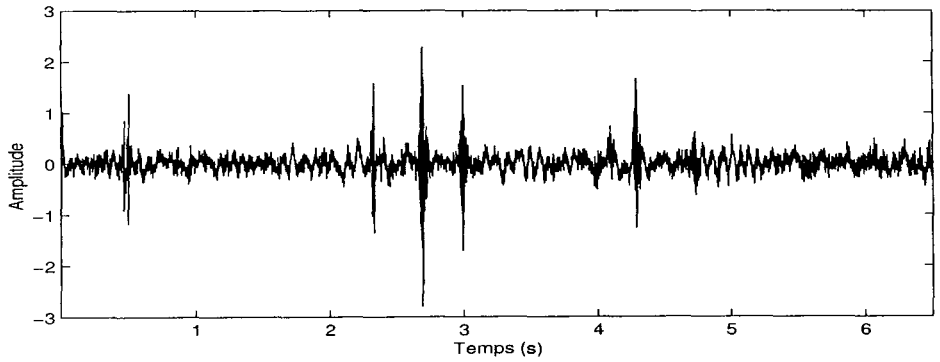


FIG. 2.1 – Exemple de signal sonore abdominal d'une durée de 6,5 secondes, enregistré avec la chaîne d'instrumentation présentée dans ce chapitre.

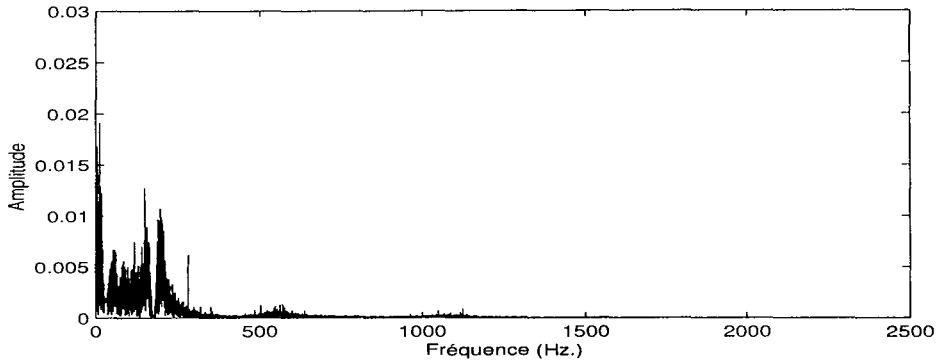


FIG. 2.2 – Spectre du signal mesuré présenté dans la Fig. 2.1 (f_e = 5000 Hz.).

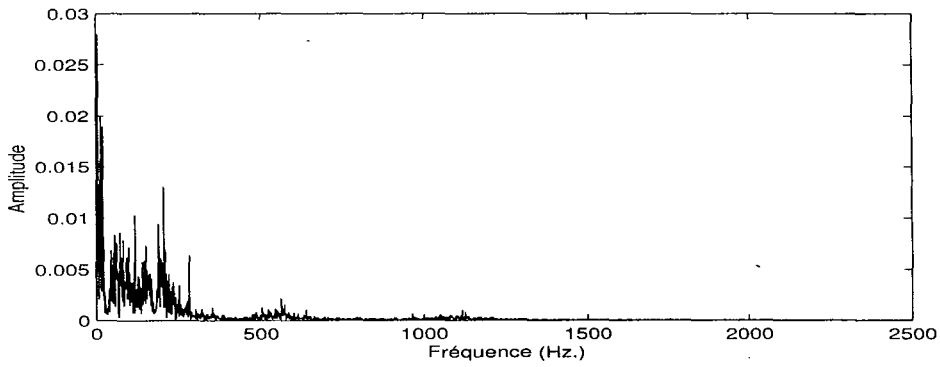


FIG. 2.3 – Spectre d'une zone de "silence" entre deux sons abdominaux réels (f_e = 5000 Hz.).

2. Quant aux caractéristiques des sons individuels, la durée est la grandeur la plus utilisée, parce qu'elle facilite la segmentation, même avec des moyens visuels ou auditifs. Cependant, les valeurs données par la bibliographie sont aussi très variables : de 20 ms. à 5 s., selon Garner et Ehrenreich [48], mais plutôt inférieures à 100 ms. [34, 84, 110, 116]. L'intensité sonore (ou des mesures liées à l'intensité, comme l'amplitude ou l'énergie) est aussi utilisée pour décrire les sons abdominaux (par ex. dans [110]), mais plus rarement, car les variations inter-individuelles sont importantes [48]. Enfin, la plupart des auteurs parlent de fréquences maximales des sons abdominaux inférieures à 1000-1500 Hz. [15, 32, 74, 84, 116, 126], même si d'autres valeurs sont mentionnées dans [34] (pour qui les fréquences maximales observées atteignent 2500 Hz., avec l'essentiel de l'énergie concentrée en dessous de 1500 Hz.) et [48] (5000 Hz.). La fréquence principale des sons abdominaux est généralement supérieure aux fréquences des sons cardiaques et pulmonaires (supérieure à 100 Hz. selon [15]) – dans la littérature on propose même un filtrage passe-haut à 80 Hz. pour éliminer leur influence [116].

Cette description du signal sonore abdominal est confirmée par nos observations. Toutes ces caractéristiques permettent d'affirmer que le signal est constitué d'une succession éparse de sons abdominaux de type impulsionnel (des transitoires), non-stationnaires (un exemple de signal expérimental est présenté figure 2.1). Ses caractéristiques temporelles varient en fonction du temps, de la localisation du son et de l'état du patient. Son contenu fréquentiel est limité : seulement environ 0,5% de l'énergie du signal est situé au-delà de 1000 Hz., et ce pourcentage monte à environ 2% pour la plage de fréquences [500 - 2500 Hz.], l'essentiel de l'énergie étant concentrée entre 100 et 500 Hz. (voir figure 2.2).

Les parties du signal qui séparent les sons, qu'on appelle dans la bibliographie "périodes de silence", ne sont pas en réalité complètement silencieuses. Le bruit dû aux effets acoustiques de cavité du stéthoscope et aux signaux sonores basse fréquence de l'organisme (respiration, flux sanguin) qui se superpose au signal informatif est un constituant important des phonoentérogrammes et doit être pris en compte dans tout traitement ultérieur. Sa puissance n'est pas négligeable par rapport à l'énergie totale du signal mesuré (signal informatif + bruit) : par exemple, pour le signal présenté figure 2.1, le rapport est de 1/4. Son contenu fréquentiel est quasiment identique à celui du signal (voir figure 2.3) : il n'est pas blanc et ne peut pas être éliminé par simple filtrage.

2.2 Protocoles de mesure

Les protocoles de mesure présentés dans la bibliographie varient d'un auteur à l'autre, en fonction des objectifs mais aussi des moyens d'acquisition et de traitement. Les durées d'enregistrement minimales varient entre une minute [15, 54, 76], deux minutes [30, 31] et cinq minutes [97]. Il est nécessaire de noter que l'objectif des auteurs de [54] est la validation d'un algorithme de débruitage, et celui de [97] vise à estimer l'effet de la palpation abdominale sur la génération des sons (les autres travaux cités étudient la physiologie abdominale normale ou pathologique). D'autres protocoles d'auscultation, dont l'objectif premier est l'étude physiologique ou le diagnostic médical, proposent des durées s'étalant de 10 minutes [5, 110], 15 minutes [126], 20 minutes [98] etc., à plusieurs heures, distribuées sur une journée complète [7, 18, 48, 122].

La standardisation des repas est courante [7, 30, 31, 32, 116, 119] mais pas obligatoire [15, 121].

En ce qui concerne la position des capteurs, on retrouve des localisations communes dans la quasi-totalité des publications : en région épi-gastrique, en région péri-ombilicale, et parfois dans le cadran inférieur gauche et dans le cadran inférieur droit. L'auscultation simultanée proposée

par Craine [32] respecte cette distribution des microphones. Dans l'article de Garner et Ehrenreich [48], deux capteurs supplémentaires sont placés au niveau des deux angles droits faits par le côlon, au début et à la fin de sa partie transverse (horizontale). Les capteurs (microphones ou stéthoscopes) sont attachés à l'abdomen (collés ou maintenus par des bandes élastiques), qui est parfois rasé afin de minimiser les frottements.

Pour minimiser les bruits parasites, les enregistrements sont généralement faits en milieu silencieux.

2.2.1 Proposition de protocole

Notre proposition de protocole de mesure s'inspire de la description précédente et de la caractérisation des signaux abdominaux.

L'activité gastro-intestinale dépend du contenu du tube digestif, de la phase du cycle digestif, de l'état de santé du patient et des éventuelles interventions chirurgicales ou médicamenteuses. Nous avons réalisé tous nos travaux sur des phonoentérogrammes des volontaires sains, qui n'étaient pas sous aucun traitement médicamenteux. Afin d'assurer un maximum de reproductibilité, tous les enregistrements ont été commencés immédiatement après un repas standardisé (petit-déjeuner).

Une autre caractéristique importante est la variation essentiellement lente de l'activité digestive. Elle se déroule sur plusieurs heures, donc toute recherche d'informations doit tenir compte de cette caractéristique : les enregistrements doivent être réalisés sur des longues durées ou à différents intervalles temporels, répartis sur la durée. Nous avons recueilli des signaux durant environ 168 minutes (à une fréquence d'échantillonnage de 5000 Hz., cette durée correspond à 12 séquences successives à 2^{22} points), afin de couvrir la phase postprandiale et au moins une partie de la phase inter-digestive.

Les contractions génératrices de sons sont produites par plusieurs organes du tube digestif, localisés aux différents endroits dans la cavité abdominale. Nous proposons un ensemble de six capteurs, placés sur la surface de la paroi abdominale (voir figure 2.4) en respectant les suggestions de Garner et Ehrenreich [48] (capteurs 1–3, 5 et 6) et Craine *et al.* [32] (1, 3 et 5). Nous avons rajouté en plus le capteur numéro 4, sous-ombilical, qui nous permet ainsi de circonscrire complètement l'aire d'intérêt.

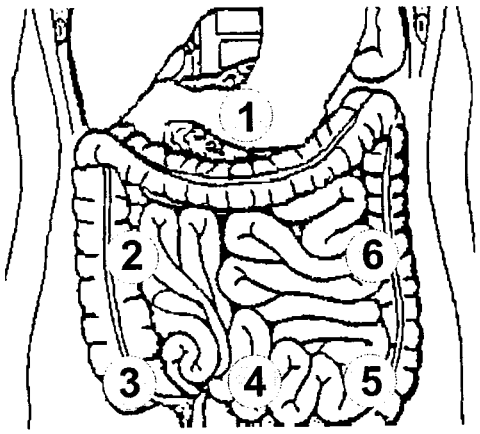


FIG. 2.4 – Placement des capteurs par rapport à l'anatomie abdominale.

Enfin, pour minimiser le nombre d'artéfacts, les volontaires ont été installés confortablement, en position demi-allongée, dans une salle silencieuse. Les stéthoscopes ont été maintenus en place

à l'aide d'une bande élastique, censée à assurer une pression d'appui relativement constante sans gêner les mouvements respiratoires et sans provoquer des frottements avec les têtes de stéthoscope ou entre celles-ci et la peau.

2.3 Chaîne d'instrumentation

L'étude bibliographique permet d'observer que l'instrumentation utilisée dépend à la fois de la description du signal selon chaque auteur et de l'objectif poursuivi (recherche physiologique, aide au diagnostic, surveillance post-opératoire ou évaluation des effets des médicaments).

Les capteurs utilisés dans la quasi-totalité des cas sont des microphones avec une bande passante suffisamment large pour couvrir le spectre des sons abdominaux. Craine *et al.* [30, 31, 32] utilisent des stéthoscopes électroniques commerciaux pour l'acquisition des sons abdominaux, alors que Garner et Ehrenreich [48] adaptent des microphones aux têtes de stéthoscope classiques. Les réponses fréquentielles présentées figure 2.5 permettent de comparer, à titre d'exemple, les stéthoscopes acoustiques classiques et leurs homologues électroniques commerciaux.

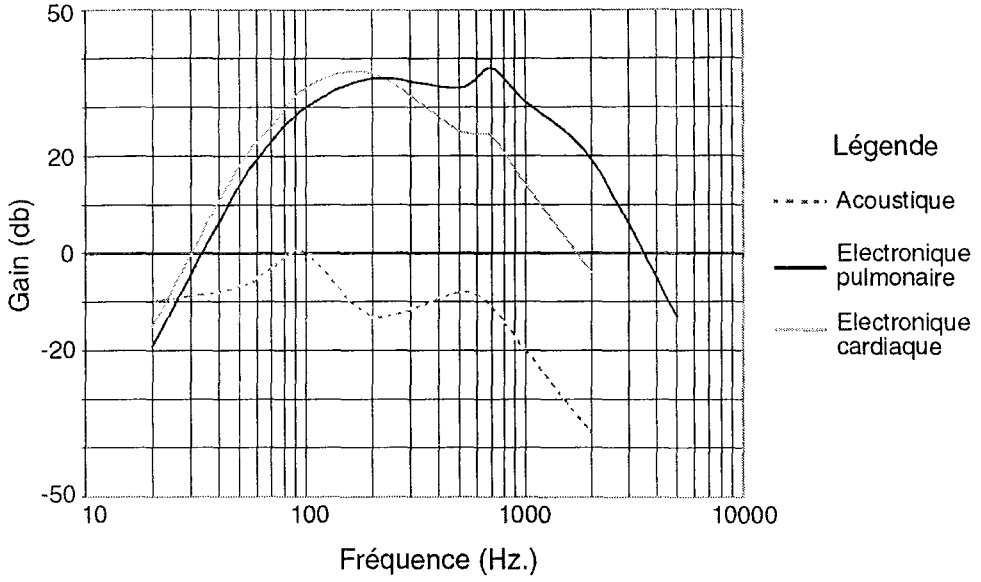


FIG. 2.5 – Exemple de réponses fréquentielles des stéthoscopes acoustiques et électroniques (les courbes présentées font partie de la documentation technique du stéthoscope électronique *E-Scope* produit par *Cardionics Inc.* `http://www.cardionics.com`, qui a deux modes de fonctionnement : cardiaque et pulmonaire).

Le traitement ultérieur est parfois analogique (integrateurs d'énergie [41], bancs de filtres [121]) et, souvent, l'interprétation des signaux est visuelle ou auditive.

Les traitements numériques proposés récemment [15, 30, 31, 32, 54, 116] sont précédés d'un conditionnement classique (amplification et filtrage anti-repliement) et sont implémentés sur des ordinateurs équipés de cartes d'acquisition et de logiciels de traitement. Évidemment, les études sur la rythmicité, ainsi que sur les phénomènes de longue durée comme le Complexe Migrant Moteur (CMM), nécessitent des enregistrements qui peuvent aller jusqu'à plusieurs heures et donc des capacités de stockage importantes.

Ces considérations, ainsi que la description précédente du signal, nous conduisent vers le choix d'une chaîne d'instrumentation classique, adaptée aux fréquences et à la dynamique du signal et à la nécessité d'enregistrements de longue durée. Le schéma général est présenté figure 2.6.

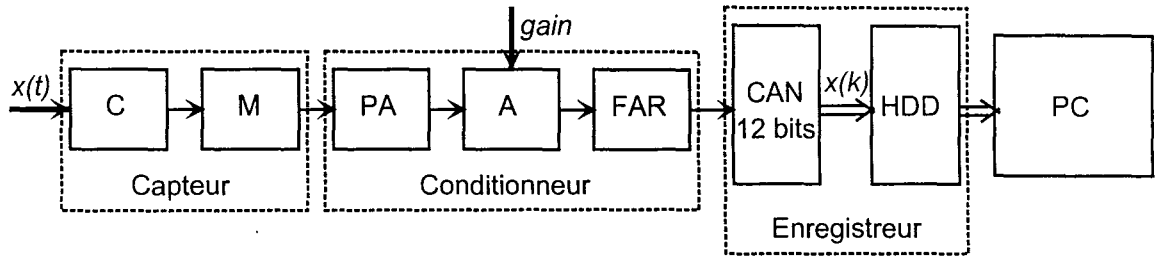


FIG. 2.6 – Schéma général de la chaîne d'instrumentation : C = tête de stéthoscope, M = microphone, PA = pré-amplificateur, A = amplificateur à gain réglable, FAR = filtre anti-repliement, CAN = convertisseur analogique-numérique, HDD = disque dur de stockage, PC = ordinateur de traitement. Le CAN et le HDD sont regroupés dans un enregistreur numérique portable industriel[9].

Les éléments de la chaîne d'instrumentation (Fig. 2.6) sont classiques, à l'exception du capteur, qui sera décrit dans une section suivante (2.3.1). Nous décrivons ici brièvement les autres parties : le conditionnement par filtrage et amplification, suivi par la numérisation et le stockage des signaux enregistrés.

2.3.1 Capteurs

Concernant le choix des capteurs, nous avons opté pour l'approche de Garner et Ehrenreich [48], qui montre que ce type d'instrumentation est performant : nous utilisons des têtes de stéthoscope classiques auxquelles sont adaptés des microphones à électret. La justification principale de ce choix est de faciliter la validation médicale, surtout pour la partie classification des sons, car l'apprentissage des médecins est fait à travers un stéthoscope acoustique classique qui possède sa propre fonction de transfert. Les signaux enregistrés seront donc semblables à ceux qu'aurait pu entendre un médecin en utilisant directement un stéthoscope qui lui est familier.

Têtes de stéthoscope. Les industriels proposent plusieurs types de têtes de stéthoscope. Nous avons choisi d'en tester deux :

- le stéthoscope standard Colson BK3001 ;
- le stéthoscope de haute qualité Littmann–3M Classic *II* S.E.

Microphone et pré-amplificateur. Le microphone est un capteur standard à électret (Philips SBC ME600), commercialisé avec un pré-amplificateur. La bande passante constructeur de l'ensemble et de 50 Hz. à 18 kHz., la tension de sortie évolue de ± 10 mV., son impédance est de 1 kΩ.

2.3.2 Conditionnement et numérisation

Comme le contenu fréquentiel des signaux sonores abdominaux est de bande limitée (d'environ 80 Hz. à 1000 Hz.), nous avons choisi une fréquence d'échantillonnage de 5000 Hz., qui permet une bonne fidélité des enregistrements tout en gardant une taille raisonnable des phonoentérogrammes numériques (2^{21} échantillons pour 168 minutes). Le signal fourni par le capteur doit être conditionné pour être adapté au dernier élément de la chaîne d'instrumentation, le convertisseur analogique-numérique (CAN). Il est donc amplifié, pour assurer la compatibilité entre son

[9]L'enregistreur numérique Nicolet Vision Portable Data Acquisition System est fabriqué par Nicolet Technologies `http://www.niti.com`.

amplitude et la plage d'entrée du CAN, et filtré, pour empêcher le repliement du spectre dû à l'échantillonnage.

Dans un souci d'assurer un rapport signal sur bruit maximal, nous avons choisi des composants électroniques de grande précision. Nous nous proposons de faire une description fonctionnelle de ce traitement analogique.

Amplificateur. L'amplificateur est construit à partir d'un amplificateur d'instrumentation particulièrement préconisé pour des applications physiologiques comme l'ECG, l'EEG et l'EMG (Burr-Brown INA126). L'impédance d'entrée très élevée de ce composant assure une bonne adaptation aux éléments précédents. Il est configuré pour fournir un gain de 100 dans la bande qui nous intéresse.

Filtrage anti-repliement Le filtrage est réalisé par deux filtres de Butterworth de second ordre en cascade, respectivement un filtre passe-haut à 60 Hz (-3 dB) et un filtre passe-bas à 2250 Hz. L'ensemble possède un gain de 10.

Numérisation et stockage. A ce stade, le signal a une amplitude maximale de ± 10 V. Il est injecté dans un système d'acquisition numérique transportable (Nicolet Vision 2.0), qui l'échantillonne à 5 kHz., sur 16 bits. Ce système assure l'acquisition parallèle de 8 voies différentielles. La capacité de stockage de l'enregistreur est de 9GB.

2.3.3 Chaîne d'instrumentation multi-voies

La description précédente prend en compte une chaîne d'instrumentation composée d'une seule voie d'acquisition, mais le traitement multi-voies s'impose dans notre application. Celui-ci ne pose pas de problème supplémentaire. En fait, il s'agit juste d'une association de plusieurs chaînes d'instrumentation identiques (figure 2.6) indépendants. Le développement que nous présentons dans la suite de ce chapitre (et dans le chapitre suivant, dédié au pré-traitement) est donc consacré à une seule voie, sachant qu'il est applicable sans modifications aux autres.

2.4 Étude de la réponse fréquentielle

Nous avons étudié la fonction de transfert de la chaîne d'instrumentation choisie, pour nous assurer de son adéquation au signal, notamment en fonction de la pression d'appui de la tête de stéthoscope sur la peau : dans la bibliographie, on ne précise pas la pression d'appui des capteurs sur l'abdomen et on ne discute pas de son éventuelle influence sur le signal recueilli. L'objectif de cette section n'est pas une modélisation précise du système, mais la vérification des éventuelles modifications en fonction des conditions d'utilisation.

Cette section continue par un rappel du principe d'identification non-paramétrique, présenté pour le cas discret – les signaux intervenant dans la modélisation sont numériques, échantillonnés avec une période T. Ensuite, nous présentons le protocole de mesure que nous avons mis au point pour l'identification et les résultats obtenus.

2.4.1 Méthode

Notre objectif n'est pas une identification paramétrique du capteur, mais l'étude de sa réponse fréquentielle, et plus précisément de son module, qui va donner le gain de la fonction de transfert du stéthoscope. Les résultats obtenus pour différentes pressions d'appui seront ensuite comparés.

En supposant les perturbations nulles, le modèle simplifié du système H à modéliser est donc présenté figure 2.7 (x étant le signal d'entrée et y celui de sortie)[10]. Les parties constitutives du système, c'est-à-dire la tête de stéthoscope, le microphone, et le conditionneur (amplificateur + filtre), sont représentées respectivement par H_C, H_M et H_A.

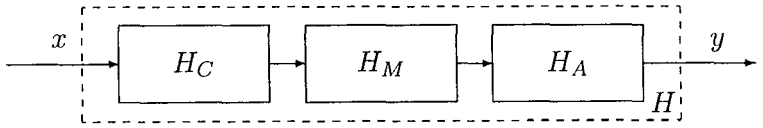


FIG. 2.7 – Modèle d'identification simplifié.

L'hypothèse de base est la linéarité et l'invariance des systèmes à caractériser. Dans ce cas, la fonction de transfert $H(i\omega)$ est définie à l'aide des transformées de Fourier discrètes $X(i\omega)$ et $Y(i\omega)$ de l'entrée et de la sortie :

$$Y(i\omega) = H(i\omega)X(i\omega), \tag{2.1}$$

avec $X(i\omega) = \sum_{k=-\infty}^{\infty} x(k)e^{-i\omega k}$ et $Y(i\omega)$ définie de manière similaire, pour $k \in \mathbb{Z}$.

Le module de la fonction de transfert du système s'écrit :

$$|H(i\omega)|^2 = \frac{\Phi_y(\omega)}{\Phi_x(\omega)}, \tag{2.2}$$

où

$$\Phi_x(\omega) = |X(i\omega)|^2 \tag{2.3}$$

désigne le spectre d'énergie du signal x, le spectre de la sortie y étant défini de manière similaire.

En réalité, on ne peut pas *calculer* les spectres qui interviennent dans les expressions antérieures des fonctions de transfert, car ni les transformées de Fourier de l'entrée x et de la sortie y, ni celles des fonctions d'auto- et intercorrélation ne sont calculables de $-\infty$ à $+\infty$. De plus, on ne peut pas obtenir leur forme analytique sans connaître les expressions de x et y. Par conséquent, on a recours à des *estimations* de ces spectres.

En suivant le développement proposé par Ljung [70], nous considérons dans les paragraphes suivants la période d'échantillonnage normalisée $T = 1$ (le passage à la bonne échelle temporelle est immédiat) et présentons plusieurs méthodes d'estimation du spectre de l'entrée $\Phi_x(\omega)$ pour la suite discrète $x(k)$, $k = 1 \ldots N$:

- le périodogramme, qui se base directement sur la définition du spectre (2.3) :

$$\hat{\Phi}_x^N(\omega) = \frac{1}{N}|X_N(i\omega)|^2, \quad \text{avec} \quad X_N(i\omega) = \sum_{k=1}^{N} x(k)e^{-i\omega k}.$$

  Cette méthode n'est pas exploitable en pratique, car les résultats sont très irréguliers (bruités) ;
- une amélioration du périodogramme est possible par la méthode de Welch, qui préconise un moyennage, c'est-à-dire le découpage du signal $x(k), k = 1 \ldots N$ en plusieurs ($j = 1 \ldots Q$)

[10] Un développement similaire à celui présenté dans cette section est possible en considérant un bruit de mesure non nul mais décorrélé de l'entrée (voir [70]).

segments de longueur M et le calcul de son spectre comme une moyenne des différents périodogrammes $\hat{\hat{\Phi}}_{x,j}^{M}(\omega)$ des segments :

$$\hat{\Phi}_x^N(\omega) = \frac{1}{Q}\sum_{j=1}^{Q}\hat{\hat{\Phi}}_{x,j}^{M}(\omega) ;$$

- la méthode de Blackman-Tukey propose un filtrage (un lissage) du périodogramme :

$$\hat{\Phi}_x^N(\omega) = \int_{-\pi}^{\pi} W_\gamma(\omega-\xi)\hat{\hat{\Phi}}_x^N(\xi)d\xi.$$

$W_\gamma(\omega)$ étant une fenêtre centrée autour de $\omega = 0$, d'intégrale unitaire $\int_{-\pi}^{\pi} W_\gamma(\omega)d\omega = 1$ et caractérisée par un paramètre γ qui détermine sa largeur.

En pratique, la méthode de Blackman-Tukey ne s'applique pas directement sur les spectres estimés par le périodogramme, mais dans le domaine temporel. Les spectres sont alors calculés à l'aide de la fonction d'autocorrélation des signaux, et l'estimation lissée de $\Phi_x(\omega)$ s'écrit [70] :

$$\hat{\Phi}_x^N(\omega) = \sum_{k=-\gamma}^{\gamma} w_\gamma(k)\hat{R}_x^N(k)e^{-i\omega k}, \tag{2.4}$$

où

$$w_\gamma(k) = \int_{-\pi}^{\pi} W_\gamma(\xi)e^{i\xi k}d\xi$$

désigne la forme temporelle de la fenêtre[11] $W_\gamma(\xi)$, et

$$\hat{R}_x^N(k) = \frac{1}{N}\sum_{t=1}^{N} x(t+k)x(t).$$

Après l'estimation similaire du spectre de la sortie $\hat{\Phi}_y^N(\omega)$, le gain de la fonction de transfert estimée du système est obtenu par :

$$|\hat{H}(i\omega)| = \sqrt{\frac{\hat{\Phi}_y^N(\omega)}{\hat{\Phi}_x^N(\omega)}}. \tag{2.5}$$

Nous avons choisi cette méthode pour identifier notre système et ses éléments[12]. L'implémentation informatique utilise l'algorithme SPA (*spectral analysis*) proposé dans Matlab.

[11] En pratique, la fenêtre choisie peut être de type Hamming, Hanning etc. La taille de la fenêtre est inversement proportionnelle à la résolution fréquentielle souhaitée.

[12] Une autre méthode utilise l'estimation du spectre croisé $\hat{\Phi}_{yx}^N(\omega)$. Dans ce cas, la fonction de transfert estimée du système est :

$$\hat{H}(i\omega) = \frac{\hat{\Phi}_{yx}^N(\omega)}{\hat{\Phi}_x^N(\omega)}. \tag{2.6}$$

Cette méthode est applicable aussi quand la sortie est perturbée par un bruit de mesure additif.

2.4.2 Protocole

Notre objectif principal n'est pas la modélisation paramétrique du système. Cependant, nous devons nous assurer que le capteur utilisé est :

- approprié, donc qu'il ne change pas la nature des signaux et que ces signaux restent cohérents avec les mesures présentées dans la littérature et interprétables par les médecins. Autrement dit, la fonction de transfert (la réponse fréquentielle) de notre capteur doit être adaptée aux signaux qui nous intéressent ;
- robuste, c'est-à-dire qu'il ne doit pas être trop sensible aux changements en fonction des conditions d'utilisation. Dans notre cas, il s'agit de la pression que la tête de stéthoscope exerce sur la peau, pression qui peut varier avec les stéthoscopes, avec les patients ou qui peut se modifier au cours du temps ou d'un enregistrement à un autre.

Ces objectifs peuvent être atteints par l'étude de l'estimation de la réponse fréquentielle.

Cette étude nécessite la définition d'un protocole de mesure, censé déterminer les signaux d'entrée utilisés, les éléments du système à modéliser et l'instrumentation nécessaire. En plus, le protocole doit prendre en compte la nature des perturbations et éventuellement les minimiser. Nous présentons dans cette section le protocole que nous avons utilisé. Sa mise en œuvre a été faite en collaboration avec l'Institut National de Recherche et de Sécurité (INRS), qui nous a fourni l'instrumentation et nous a assuré les conditions nécessaires[13].

Instrumentation et conditions de test

Comme la bande fréquentielle du signal d'intérêt dans notre application a sa borne supérieure à 1000–1500 Hz., nous avons échantillonné les signaux recueillis à 5000 Hz. La chaîne d'instrumentation, qui sera utilisée ensuite pour les enregistrements des sons abdominaux expérimentaux, est présentée dans la section 2.3.2.

Pour pouvoir identifier un système, on doit avoir une bonne connaissance du signal d'entrée. De plus, pour pouvoir utiliser directement l'estimation du gain de la fonction de transfert (éq. (2.5)), il est nécessaire de minimiser les perturbations (ou de s'assurer qu'elles sont indépendantes de l'entrée). La mise en oeuvre de notre protocole de test passe donc par le choix de l'environnement et de la source sonore.

Environnement. Les enregistrements nécessaires ont été effectués dans une chambre anéchoïde à l'INRS. Nous avons ainsi éliminé les bruits extérieurs et les échos dus aux réflexions du signal sonore émis par la source ponctuelle. Sous ces conditions, le développement présenté dans la section 2.4.1 est adapté à l'identification de notre capteur : les perturbations sont considérées nulles et les équations (2.2) et (2.5) sont applicables.

Étude et calibrage de la source. Pour pouvoir estimer la réponse fréquentielle du système, il est souhaitable que le signal d'entrée soit stationnaire durant le protocole de mesure et qu'il contienne toutes les fréquences dans la plage qui nous intéresse (80–1000 Hz.). Nous avons donc utilisé une source S_1, considérée ponctuelle, réglée pour émettre un bruit relativement proche d'un bruit blanc dans cette bande de fréquence. Le son généré a été enregistré à l'aide d'un microphone de haute précision[14] M_0. Nous avons considéré le signal fourni par ce microphone comme une estimation fidèle $\hat{S}_1$ de la source et nous l'avons utilisé pour calculer son spectre.

[13]Nous remercions à M. T. Loyau, chercheur à l'INRS, pour ses conseils avisés et sa disponibilité.
[14]Bruel & Kjaer.

La première étape a été le calibrage du microphone M_0, ayant une fonction de transfert $H_{M_0} = G_0$ de gain inconnu mais constant dans la bande passante. Ce calibrage, nécessaire pour déterminer le gain G_0, a été effectué à l'aide d'une source sonore sinusoïdale monochromatique de précision (S_0), émettant un son continu à 1000 Hz., à une intensité acoustique constante de 94 dB.

Par définition, l'intensité acoustique en décibels est :

$$I(dB) = 10\log_{10}\frac{I}{I_a} = 10\log_{10}\frac{p^2}{p_a^2}, \tag{2.7}$$

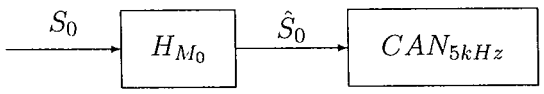
où $p_a = 2\cdot 10^{-5}$ (N/m^2) est la pression acoustique correspondant au seuil d'audibilité $I_a = const.$ et p la pression acoustique mesurée [81]. Un calcul direct de (2.7), en remplaçant les valeurs de I et de p_a, montre que le niveau sonore standard de 94 dB de la source de calibrage assure une pression acoustique d'amplitude unitaire. Comme la relation entre la pression acoustique de la source S_0 (l'entrée) et le signal fourni par le microphone $\hat{S}_0$ (la sortie) est $\hat{S}_0 = G_0 \cdot S_0$, le gain G_0 est donné par l'amplitude du signal sinusoïdal de sortie $\hat{S}_0$ (voir Fig. 2.8(a)).

La mesure directe de l'amplitude du signal de sortie nous donne :

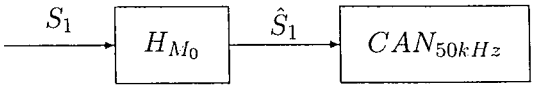
$$G_0 = 1,25. \tag{2.8}$$

Ce gain est utilisé ensuite pour normaliser la source de bruit S_1 (par multiplication du signal $\hat{S}_1$ par $H_{M_0}^{-1} = G_0^{-1}$), étape nécessaire avant une estimation correcte de l'amplitude du spectre $\hat{\Phi}_{S_1}$[15] (voir Fig. 2.8(c)).

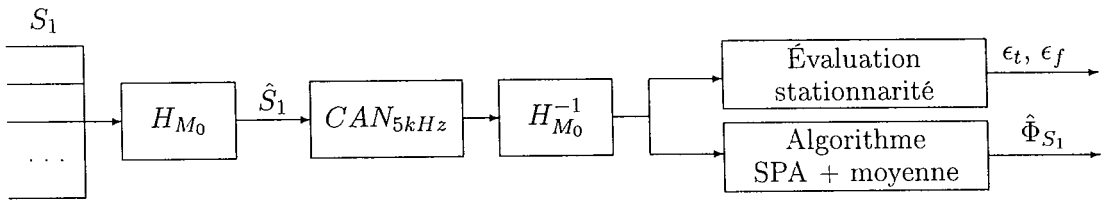
Enfin, les dernières étapes de calibrage (voir Fig. 2.8(b) et 2.8(c)) ont été le réglage du contenu fréquentiel de la source sonore S_1 utilisée comme signal d'entrée, la vérification de sa stationnarité et l'estimation de son spectre $\hat{\Phi}_{S_1}$.



(a) calibrage du microphone M_0



(b) calibrage fréquentiel de la source S_1



(c) vérification de la stationnarité (source S_1) et estimation du spectre $\hat{\Phi}_{S_1}$

FIG. 2.8 – Protocole de mesure – étapes de calibrage.

La source S_1 a été réglée pour qu'elle émette un signal contenant toutes les fréquences entre 100 et 1000 Hz., si possible avec des puissances proches. Plusieurs mesures ont été effectuées :

[15] Cette étape permet d'évaluer, à titre informatif, l'intensité acoustique de la source S_1 : après le calcul de l'enveloppe par une transformée de Hilbert et en considérant la moyenne de cette enveloppe comme l'amplitude de l'onde de pression acoustique, l'intensité calculée à l'aide de l'équation (2.7) est I_{S_1}=87 dB.

- une mesure à une fréquence d'échantillonnage de calibrage de 50 kHz., pour estimer l'énergie émise par la source au-delà de 2500 Hz. Cette estimation est nécessaire pour s'assurer que la fréquence d'échantillonnage choisie pour l'identification du système (5000 Hz.) est suffisante, vu qu'aucun filtre anti-repliement n'a été utilisé pour le signal $\hat{S}_1$ fourni par M_0 durant les mesures d'identification ;
- plusieurs mesures décalées dans le temps, pour s'assurer de la stationnarité de la source.

Nous avons estimé les spectres des signaux par la méthode de Blackman-Tukey (éq. (2.4)), après avoir normalisé en puissance la source S_1 par division du signal recueilli par G_0 (éq (2.8)). Le spectre de la source, utilisant la fréquence d'échantillonnage de calibrage : $f_{ec} = 50$ kHz., est présenté figure 2.9. Même s'il est évident que la source n'est pas blanche, toutes les fréquences d'intérêt sont présentes (voir tableau 2.1). L'énergie du signal située dans les hautes fréquences (> 2500 Hz.) est très faible ($< 0{,}1\%$ de l'énergie totale), donc une fréquence d'échantillonnage à 5000 Hz. est suffisante.

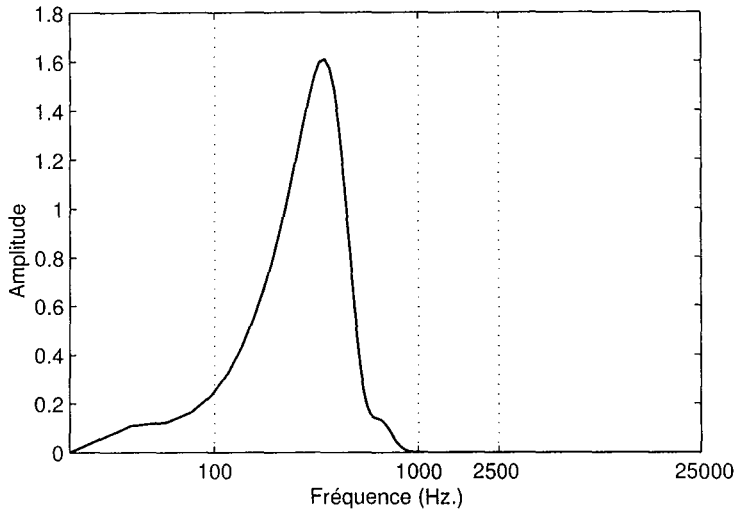


FIG. 2.9 – Le spectre de la source de bruit utilisée (S_1) (échelle logarithmique pour les fréquences). La fréquence d'échantillonnage choisie pour le calibrage est de 50 kHz., et la source est réglée pour être relativement "blanche" entre 100 et 1000 Hz.

Fréquence (Hz.)					
< 50	50 – 100	100 – 500	500 – 1000	1000 – 2500	> 2500
0,4%	2,1%	86,9%	10,4%	0,06%	0,06%

TAB. 2.1 – Répartition énergétique de la source de test S_1 par bande fréquentielle. Les valeurs doivent être interprétées comme des ordres de grandeur, car l'estimation du spectre (algorithme SPA) dépend de la résolution fréquentielle choisie.

Pour s'assurer de la stationnarité de la source, ses paramètres statistiques (écart-type et moyenne[16]), ainsi que les spectres de puissance (éq. (2.4)), ont été estimés pour plusieurs mesures. L'écart de la stationnarité statistique est quantifié par deux indicateurs que nous définissons par la suite :

- **Erreur de stationnarité temporelle ϵ_t.** Nous avons estimé les écarts-type $\hat{\sigma}_{S_1,i}$ pour $i = 1 \ldots 26$ signaux (enregistrements de la source) et calculé un écart-type moyen $\hat{\sigma}_{S_1,m}$.

[16] La moyenne est toujours nulle par la nature même du signal sonore et de l'instrumentation.

Nous avons défini l'erreur relative comme la valeur absolue de la différence normalisée entre l'écart-type estimé du signal i et l'écart-type moyen :

$$\epsilon_{t,i} = \left| \frac{\hat{\sigma}_{S_1,i} - \hat{\sigma}_{S_1,m}}{\hat{\sigma}_{S_1,m}} \right|.$$

La valeur maximale de l'erreur de stationnarité temporelle est :

$$\max_i(\epsilon_{t,i}) = 0,03 \ (3\%),$$

et sa moyenne est :

$$\hat{\epsilon}_t = \frac{1}{26} \sum_{i=1}^{26} \epsilon_{t,i} = 0,013 \ (1,3\%);$$

- **Erreur de stationnarité fréquentielle ϵ_f.** La même procédure a été appliquée pour vérifier la stationnarité des spectres estimés $\hat{\Phi}_{S_1,i}(\omega)$, pour la plage de fréquences comprise entre 80 et 1000 Hz. Cependant, étant donné qu'il s'agit de vecteurs de dimension N et non pas de valeurs scalaires comme pour les écarts-types, nous avons défini le spectre moyen $\hat{\Phi}_{S_1,m}(\omega)$ par :

$$\hat{\Phi}_{S_1,m}(\omega) = \frac{1}{26} \sum_{i=1}^{26} \hat{\Phi}_{S_1,i}(\omega), \tag{2.9}$$

et l'erreur relative du spectre du signal i par :

$$\varepsilon_{\Phi,i}(\omega) = \left| \frac{\hat{\Phi}_{S_1,i}(\omega) - \hat{\Phi}_{S_1,m}(\omega)}{\hat{\Phi}_{S_1,m}(\omega)} \right|. \tag{2.10}$$

L'erreur de stationnarité fréquentielle sera définie comme :

$$\epsilon_{f,i} = \frac{1}{N} \sum_{\omega}^{N} \varepsilon_{\Phi,i}(\omega), \tag{2.11}$$

avec $\hat{\Phi}_{S_1,m}(\omega)$ le spectre moyen de la source S_1. La valeur maximale de l'erreur de stationnarité fréquentielle est :

$$\max_i(\epsilon_{f,i}) = 0,102 \ (10,2\%),$$

et sa moyenne est :

$$\hat{\epsilon}_f = \frac{1}{26} \sum_{i=1}^{26} \epsilon_{f,i} = 0,062 \ (6,2\%);$$

L'erreur semble plus importante dans le domaine fréquentiel, mais l'essentiel de la différence est dû aux fréquences supérieures à 500 Hz., où la puissance de la source est relativement faible, les estimations moins fiables et donc le risque d'erreur relative plus grand (voir Fig. 2.10). Graphiquement, les 26 estimations des spectres $\hat{\Phi}_{S_1,i}(\omega)$ sont pratiquement superposées et se confondent avec le spectre moyen $\hat{\Phi}_{S_1,m}(\omega)$. L'allure de ce spectre moyen est présentée, en pointillé, dans la figure 2.10. Le spectre moyen $\hat{\Phi}_{S_1,m}(\omega)$ sera utilisé ensuite comme le spectre du signal d'entrée $\hat{\Phi}_x^N(\omega)$ pour estimer le module de la fonction de transfert (éq. (2.5)).

Les résultats présentés montrent que cette source est appropriée et utilisable pour les étapes suivantes de l'identification.

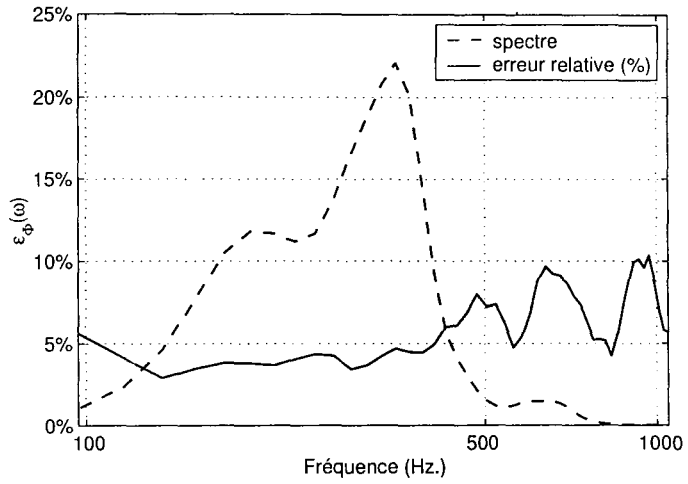
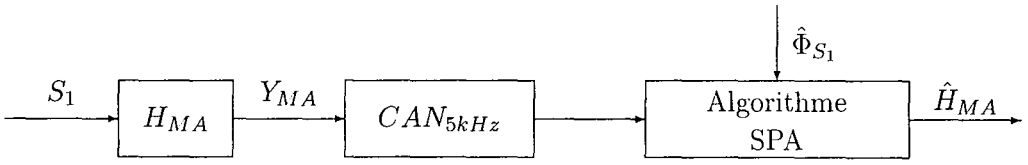


FIG. 2.10 – Variation de l'erreur de stationnarité fréquentielle $\epsilon_\Phi(\omega)$. L'erreur est superposée au spectre estimé moyen $\hat{\Phi}_{S_1,m}(\omega)$ de la source S_1. La fréquence d'échantillonnage choisie pour les mesures est de 5000 Hz.

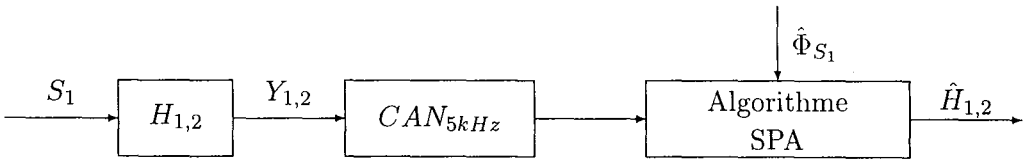
Déroulement du protocole

Fonctions de transfert. Nous avons estimé la fonction de transfert H du système (Fig. 2.7) en utilisant l'équation (2.5). Cependant, le conditionneur (voir Fig. 2.6) est un élément fixe de notre chaîne d'instrumentation. Pour simplifier la notation, nous le considérons comme partie intégrante des éléments précédents (tête de stéthoscope et/ou microphone). Ainsi, nous allons noter par la suite $H_{MA} = H_M \cdot H_A$, où H_M et H_A désignent les fonctions de transfert du microphone et du conditionneur (voir Fig. 2.7).

Un composant essentiel de notre capteur est le microphone M attaché à la tête de stéthoscope. Pour s'assurer que sa bande passante est compatible avec les signaux d'intérêt, nous avons estimé le spectre du signal fourni (après conditionneur) Y_{MA} et la fonction de transfert H_{MA} (voir Fig. 2.11(a)), conformément à l'équation (2.5).



(a) estimation de la fonction de transfert H_{MA}



(b) estimation des fonctions de transfert $\hat{H}_{1,2}$

FIG. 2.11 – Protocole de mesure - étapes d'estimation des fonctions de transfert.

Ensuite, la fonction de transfert du système H a été estimée pour deux types de têtes de stéthoscope :

- le stéthoscope Colson BK3001, noté par la suite C_1 ;

– le stéthoscope Littmann–3M Classic *II* S.E., noté C_2.

Les deux fonctions de transfert, estimées selon (2.5), sont notées par la suite $H_1 = H_{C_1} \cdot H_{MA}$ et $H_2 = H_{C_2} \cdot H_{MA}$ (voir Fig. 2.11(b)).

Influence de la pression. Enfin, pour évaluer l'influence de la pression d'appui sur la peau, nous avons enregistré les signaux de sortie et estimé les spectres et les fonctions de transfert $H_1^{(p)}$ et $H_2^{(p)}$ pour différentes pressions (figure 2.13). Le dispositif mécanique utilisé nous a permis d'augmenter graduellement la pression d'appui à l'aide des poids étalonnés de 100 à 400 grammes, avec un pas de 50 grammes. Le schéma simplifié de ce dispositif est présenté figure 2.12, et le diagramme du protocole de mesure correspondant figure 2.13. On remarque l'existence d'un élément supplémentaire – la membrane élastique qui supporte la tête de stéthoscope. Sa fonction de transfert H_S modifie la fonction de transfert totale du système et est, à son tour, influencée par la pression d'appui.

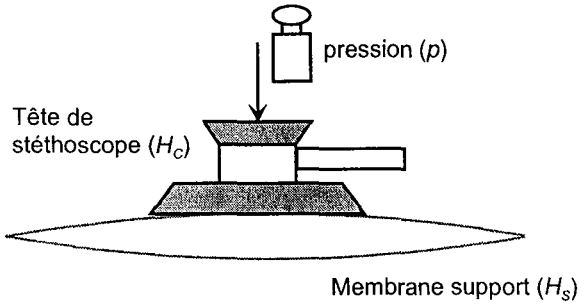


FIG. 2.12 – Schéma simplifié du dispositif pour l'étude de l'influence de la pression d'appui sur la fonction de transfert du système d'acquisition des sons.

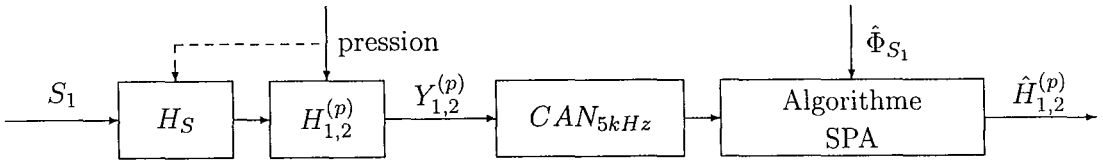


FIG. 2.13 – Protocole de mesure – estimations des fonctions de transfert $\hat{H}_{1,2}^{(p)}$ en fonction de la pression d'appui.

2.4.3 Résultats

Réponses fréquentielles

Les premières étapes consistent en l'estimation des réponses fréquentielles du système d'acquisition des sons en conditions "idéales" de fonctionnement, sans pression d'appui susceptible de les modifier.

L'estimation de la réponse fréquentielle du microphone M et de la chaîne d'instrumentation utilisés (H_{MA} dans la figure 2.11(a)) montre, sans surprise, qu'ils sont adaptés à l'acquisition des sons abdominaux (voir Fig. 2.14).

La figure 2.15 permet l'étude comparative des fonctions de transfert des deux versions de notre capteur (la chaîne d'instrumentation incluse). Les performances sont différentes, principalement dans les basses fréquences. Cependant, en comparant ces réponses fréquentielles à celles

présentées figure 2.5, on observe que les deux variantes de notre chaîne d'instrumentation offrent une meilleure amplification des sons par rapport aux stéthoscopes acoustiques. L'atténuation plus importante de la tête de stéthoscope Colson (H_1) dans les fréquences inférieures à 500 Hz. est sans doute due à ses caractéristiques mécaniques.

Une autre comparaison intéressante peut être faite entre les fonctions de transfert de notre système d'acquisition et celles des stéthoscopes électroniques commerciaux, présentées figure 2.5. En effet, l'amplification de notre capteur est moins importante dans les basses fréquences (100 – 500 Hz.). Deux explications nous semblent possibles :

- la plus faible amplification peut être due aux caractéristiques mécaniques des têtes de stéthoscope. Cependant, la réponse fréquentielle d'un stéthoscope acoustique, telle qu'elle est présentée figure 2.5, ne semble pas indiquer une atténuation importante dans cette plage de fréquence. Une hypothèse possible, qui reste à vérifier dans des travaux futurs plus axés sur l'instrumentation, est que la réponse fréquentielle d'un stéthoscope acoustique *complet* dépend également des tubes en caoutchouc qui conduisent le son entre la tête et les embouts auriculaires ;
- la courbe plus plate des stéthoscopes électroniques commerciaux, équivalente à un gain moins dépendant de la fréquence et plus important en dessous de 500 Hz., peut être obtenue grâce au traitement du signal intégré (amplification plus forte des basses fréquences).

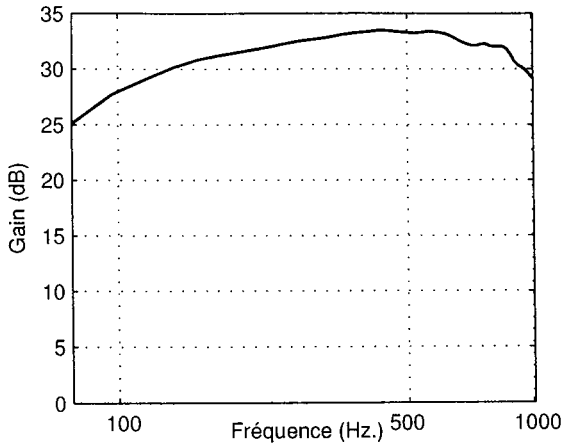


FIG. 2.14 – Réponse fréquentielle du microphone M utilisé dans notre capteur.

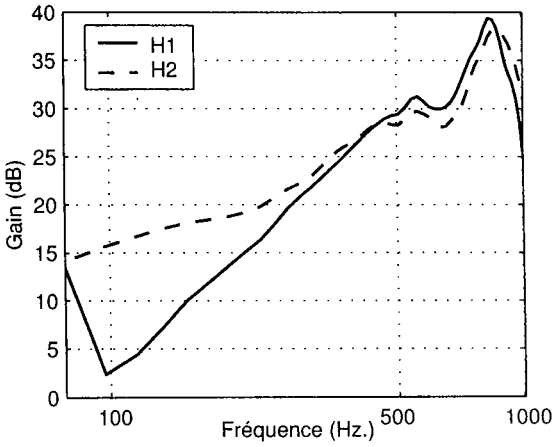


FIG. 2.15 – Réponse fréquentielle des systèmes identifiés (stéthoscope Colson = H_1, stéthoscope Littmann = H_2).

Influence de la pression

Durant la seconde étape d'identification, nous introduisons un paramètre supplémentaire : la pression d'appui sur la peau des têtes de stéthoscopes. Cette pression est susceptible de modifier la courbure des membranes et donc les caractéristiques mécano-acoustiques des cavités des têtes de stéthoscope et leurs fonctions de transfert H_C.

A l'aide du dispositif mécanique présenté dans la figure 2.12, qui introduit sa propre fonction de transfert dans la chaîne d'instrumentation, nous avons fait varier la pression d'appui entre 100 et 400 grammes, avec un pas de 50 grammes. Nous avons donc estimé les fonctions de transfert modifiées pour les p_i pressions d'appui ($i = 1 \ldots 7$). Les estimations superposées des

gains des différentes réponses fréquentielles des systèmes étudiés[17] $\hat{H}_1^{(p_i)}$ et $\hat{H}_2^{(p_i)}$ sont présentées respectivement sur les figures 2.16 et 2.17. On observe que, pour chacun des capteurs, les fonctions de transfert ont la même allure et que la pression d'appui ne semble pas influencer d'une manière significative l'acquisition des signaux.

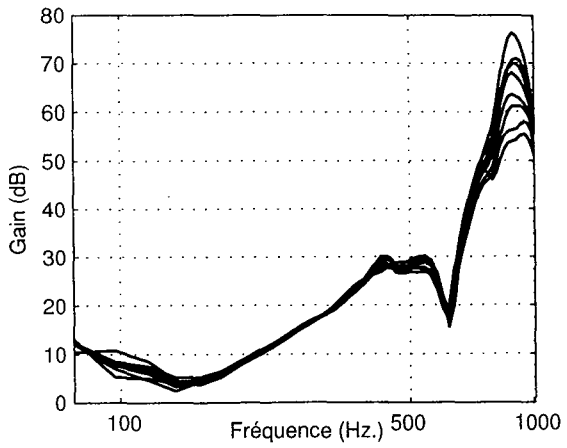


FIG. 2.16 – Superposition des réponses fréquentielles du système équipé du capteur C_1 (Colson), estimées pour différentes pressions d'appui.

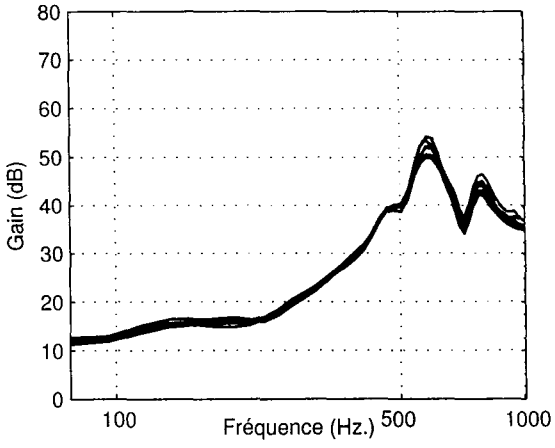


FIG. 2.17 – Superposition des réponses fréquentielles du système équipé du capteur C_2 (Littmann), estimées pour différentes pressions d'appui.

Pour quantifier l'influence de cette pression, nous avons recours à un indicateur similaire à l'erreur de stationnarité fréquentielle introduite précédemment (éq. (2.11)). Ainsi, pour le capteur C_1 (Colson) et pour une plage fréquentielle $f = 80 \ldots 1000$ Hz. (et pour la pulsation ω discrétisée correspondante), nous définissons la réponse fréquentielle moyenne $\hat{H}_1^m(\omega)$ par :

$$\hat{H}_1^m(\omega) = \frac{1}{7} \sum_{i=1}^{7} \hat{H}_1^{(p_i)}(\omega), \tag{2.12}$$

et l'écart relatif d'une estimation $\hat{H}_1^{(p_i)}(\omega)$ par rapport à cette moyenne par :

$$\varepsilon_{H_1,p_i}(\omega) = \left| \frac{\hat{H}_1^{(p_i)}(\omega) - \hat{H}_1^m(\omega)}{\hat{H}_1^m(\omega)} \right|. \tag{2.13}$$

Comme l'algorithme d'estimation des spectres (SPA) calcule les valeurs de ces spectres pour des fréquences discrètes f_j, $j = 1 \ldots N$ (et les ω_j correspondants) et que donc la fonction de transfert estimée est un vecteur de dimension N dépendant de la résolution fréquentielle choisie, nous pouvons calculer l'écart moyen de l'estimation d'une fonction de transfert $\hat{H}_1^{(p_i)}$:

$$\epsilon_{1,p_i} = \frac{1}{N} \sum_{\substack{\omega_j \\ j=1..N}} \varepsilon_{H_1,p_i}(\omega_j). \tag{2.14}$$

La valeur maximale de ϵ_{1,p_i} est :

$$\max_{p_i}(\epsilon_{1,p_i}) = 0,028 \ (2,8\%),$$

[17] pour les deux têtes de stéthoscope, Colson C_1 et Littmann C_2.

et sa moyenne (pour les 7 estimations) est :

$$\hat{\epsilon}_{1,p} = \frac{1}{7}\sum_{i=1}^{7} \epsilon_{1,p_i} = 0,015\ (1,5\%).$$

D'une manière similaire, nous avons obtenu le maximum et la moyenne des écarts des estimations pour le deuxième capteur C_2 (Littmann) :

$$\max_{p_i}(\epsilon_{2,p_i}) = 0,024\ (2,4\%),$$

$$\hat{\epsilon}_{2,p} = \frac{1}{7}\sum_{i=1}^{7} \epsilon_{2,p_i} = 0,010\ (1\%).$$

Discussion

Les résultats des estimations précédentes permettent une comparaison qualitative et quantitative des deux variantes de notre système d'acquisition. La première comparaison possible est celle entre les fonctions de transfert H_1 et H_2 des systèmes sans pression d'appui. En effet, l'interprétation directe des courbes de gains présentées dans la figure 2.15 montre que l'amplification du système H_2 (tête de stéthoscope Littmann) est plus importante dans les basses fréquences, ce qui assure une meilleure acquisition des composantes graves des sons abdominaux. Cet avantage est perdu dans les hautes fréquences (supérieures à 500 Hz.), où les deux capteurs sont équivalents.

En revanche, l'interprétation des courbes présentées sur les figures 2.16 et 2.17 doit être faite avec plus de précaution. L'allure différente de ces courbes par rapport à celles de la figure 2.15 s'explique sans doute par l'introduction de la membrane de support des stéthoscopes et donc de sa fonction de transfert H_S comme élément supplémentaire dans la chaîne de mesure (voir figure 2.12). Nous avons d'ailleurs effectué plusieurs mesures en modifiant la tension élastique de la membrane de support, et ces mesures confirment l'influence importante de cette membrane sur les résultats de l'estimation des réponses fréquentielles.

De plus, comme les deux têtes de stéthoscope testées ont une construction différente, leur interaction avec la membrane de support est à son tour différente et les fonctions de transfert totales ne sont pas modifiées de la même manière. Ce constat explique probablement l'allure différente des réponses fréquentielles des deux systèmes complets (membrane de support incluse) dans les hautes fréquences (au-delà de 500 Hz.), ce qui n'était pas le cas pour les réponses sans membrane de support présentées figure 2.15.

Néanmoins, pour chacun des deux systèmes testés, les gains des fonctions de transfert ont la même allure, l'influence de la pression d'appui étant relativement peu importante. Ce constat visuel est confirmé par l'évaluation quantitative des écarts de ces gains par rapport à leur moyenne. Les valeurs calculées de ces écarts (page 40) semblent indiquer que le système équipé du capteur C_2 (tête de stéthoscope Littmann) est moins influencé par la pression d'appui. Cependant, nous ne pouvons pas affirmer avec certitude que cette différence est due à la qualité du capteur utilisé, car l'interaction tête de stéthoscope–membrane de support a aussi un rôle important[18].

[18]Pour une tension plus grande de la membrane élastique de support, les différences visibles dans les hautes fréquences entre les différentes réponses fréquentielles $H_1^{(p)}$ sont beaucoup moins importantes, ce qui justifie notre prudence d'interprétation. Les résultats exacts ne sont pas présentés ici.

2.5 Conclusion

Dans ce chapitre, nous définissons et décrivons une chaîne d'instrumentation adaptée à l'acquisition des signaux sonores abdominaux. La nature du signal étudié nous conduit vers une chaîne d'instrumentation classique, qui doit respecter le contenu fréquentiel des signaux sonores abdominaux et assurer une grande capacité de stockage.

La nécessité d'une validation médicale simple favorise le choix d'un type de capteur déjà utilisé dans le domaine de la phonoentérographie et habituellement employé en routine clinique – le stéthoscope. En revanche, la décision concernant le type de stéthoscope utilisé est moins facile. En effet, l'étude des réponses fréquentielles présentée dans ce chapitre montre qu'aucun des stéthoscopes testés ne se détache fortement des autres. Même si les stéthoscopes électroniques commerciaux ont l'avantage d'être des outils déjà validés, avec une bande passante bien adaptée aux sons physiologiques, leur prix est élevé. Le même argument est valable pour les stéthoscopes de haute qualité (Littmann). Nous avons donc choisi, pour des raisons purement économiques, le capteur C_1 (tête de stéthoscope Colson), du moins pour le développement d'un premier outil d'auscultation automatisé. Néanmoins, d'autres méthodes d'acquisition devront être étudiées par la suite.



Détection et débruitage par ondelettes

Nous avons vu dans les chapitres précédents que les signaux d'intérêt sont constitués d'une succession d'événements impulsionnels non-stationnaires, auxquels il se superpose un bruit additif coloré. Les événements sont répartis sur de longues durées et ont des fréquences principales comprises entre 100 et 1000 Hz. En analysant la problématique présentée, on peut scinder l'étude du signal en deux niveaux :

- le premier, dédié à une analyse locale, analyse la structure des événements pour extraire leurs caractéristiques fréquentielles, temporelles et énergétiques. Les événements seront caractérisés et classés ultérieurement en fonction de leur description ;
- le second étudie statistiquement les variations au sein de la population d'événements afin de tirer des informations globales.

Les deux points précédents demandent un pré-traitement permettant de détecter les événements (transitoires non-stationnaires) et éventuellement les extraire du bruit (supposé stationnaire). La détection des événements, qui est parfois abordée sous l'angle statistique du test d'hypothèses [6, 16, 63, 91], est utilisée pour la segmentation du signal. Dans la littérature [16, 69, 91, 123] on trouve plusieurs méthodes directes de segmentation dans le domaine temporel, comme les détecteurs de saut appliqués à certains paramètres du signal : saut de moyenne, de variance, voire de moments d'ordre supérieur pour la localisation des transitoires non-gaussiens. Outre le problème du seuillage, qui se pose d'ailleurs aussi dans le cas d'une segmentation sur l'amplitude de l'enveloppe du signal, ces méthodes n'éliminent pas le bruit additif superposé aux événements.

Une autre famille de détecteurs s'applique après une modélisation du signal : le bruit additif est généralement supposé blanc, gaussien et stationnaire, alors que les événements sont modélisés comme des processus auto-régressifs. La segmentation est réalisée par une détection de changement de modèle. Une fois les coefficients du modèle AR d'un événement estimés, ils peuvent être utilisés pour la classification des événements ou pour leur analyse spectrale [16, 63, 91, 115, 123]. Un passage en revue rapide de la bibliographie montre que cette approche est souvent utilisée dans le domaine bio-médical, pour la segmentation de l'électromyogramme (EMG) utérin [63], de l'électroencéphalogramme [123], mais aussi dans le diagnostic industriel [16], l'acoustique sous-marine [91] ou la segmentation de la parole [115].

La modélisation auto-régressive n'est évidemment pas la seule approche dans l'estimation du signal. Dans l'introduction de [6], Antoniadis *et al.* énumèrer[‘] plusieurs méthodes de "régression non-paramétrique" : les estimateurs linéaires, comme les fonctions de noyau (*kernel functions*), les splines et les séries orthogonales, très étudiés durant les années 1980 et 1990, auxquels se rajoutent progressivement, après 1990, les estimateurs non-linéaires par seuillage des décompositions en ondelettes[19,20], qui donnent souvent de meilleurs résultats. La décomposition en ondelettes n'est pas à proprement parler une modélisation du signal, mais la transformation de celui-ci dans un autre domaine (temps-échelle), qui permet l'étude détaillée de ses caractéristiques : par leur construction, les ondelettes sont des fonctions localisées en temps et en fréquence, elles s'apparentent donc aux représentations temps-fréquence.

La décomposition en ondelettes est donc une méthode qui répond aux exigences des deux niveaux d'analyse rappelés en début de ce chapitre, ainsi qu'aux caractéristiques du signal : elle permet à la fois le débruitage et la segmentation des transitoires et elle offre l'information conjointe en temps et en fréquence tout en gardant une charge de calcul adaptée aux données de grande taille. Les travaux sur la détection et le débruitage par ondelettes sont très nombreux et très divers (voir par exemple [118] pour un passage en revue des applications bio-médicales des ondelettes). Dans le domaine qui nous intéresse plus particulièrement, la phonoentérographie, l'analyse en ondelettes a été déjà appliquée avec succès pour l'extraction des événements non-stationnaires du bruit par Hadjileontiadis *et al.* [54].

Ce chapitre commence par une introduction relative à la théorie des décompositions en ondelettes, suivie par un bref état de l'art sur la détection des transitoires et le débruitage par ondelettes. Les détails de la transformation en ondelettes appliquée, ainsi que les développements concernant le débruitage et la segmentation, font l'objet des sections suivantes. L'algorithme de débruitage optimisé par une itération de point-fixe est une contribution méthodologique dans le domaine du débruitage par ondelettes et un des points importants que nous présentons dans cette thèse [86, 87, 89].

3.1 Analyse temps-fréquence. Ondelettes

La représentation temps-fréquence cherche à mettre en évidence les propriétés d'un signal à la fois en temps et en fréquence. Elle est située à mi-chemin entre la représentation purement temporelle type Dirac (figure 3.1) et la représentation purement fréquentielle type Fourier (figure 3.2). Cette approche offre, en même temps, des informations sur les caractéristiques temporelles et fréquentielles d'un signal (figure 3.3).

Évidemment, la richesse informationnelle obtenue par cette méthode a son prix : si, dans une représentation type Dirac on a une résolution temporelle maximale (sans avoir aucune information fréquentielle) et, dans une représentation Fourier, on a une résolution fréquentielle maximale (sans avoir aucune information temporelle), dans le plan temps-fréquence on perd cette précision, en gagnant une information conjointe [8, 43, 57]. La précision limitée des représentations temps-fréquence trouve son explication dans le principe d'incertitude de Heisenberg-Gabor [8, 43]. On montre qu'on ne peut pas avoir accès à la fréquence instantanée d'un signal, car on ne peut pas avoir en même temps une largeur de bande fréquentielle B et une localisation temporelle dans une fenêtre T arbitrairement petites. Cependant, on peut chercher des signaux proches

[19]Dans le cas le plus courant en pratique d'une décomposition orthogonale en ondelettes, cette dernière méthode peut être interprétée comme un développement de l'estimation par séries orthogonales [6].

[20]Khalil et Duchêne proposent dans [63] une méthode combinant la décomposition en ondelettes et la modélisation AR des séries des coefficients, échelle par échelle.

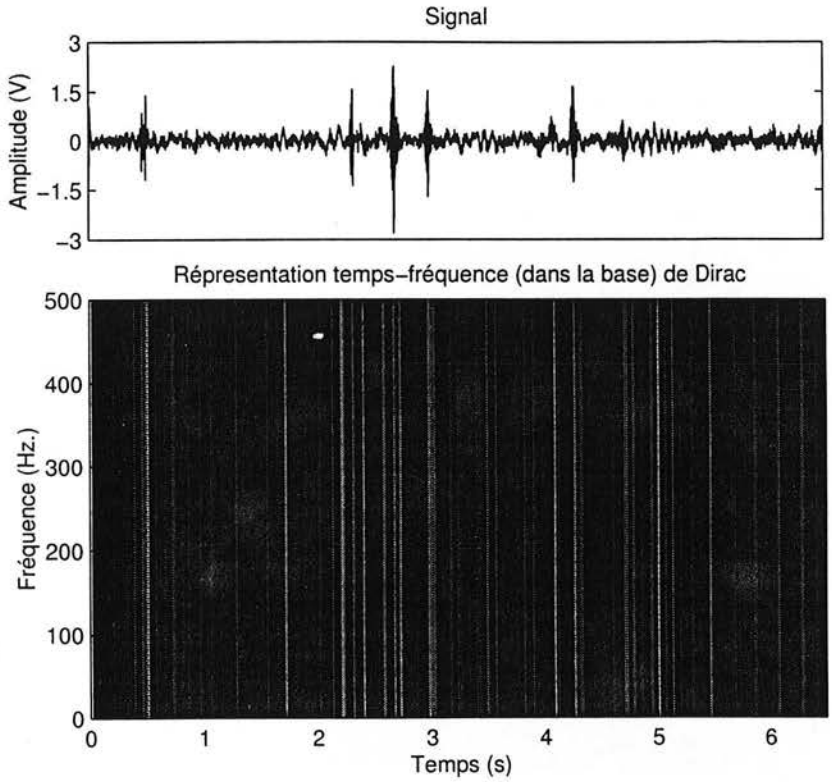


FIG. 3.1 – Représentations temporelles (Dirac) 1D et 2D

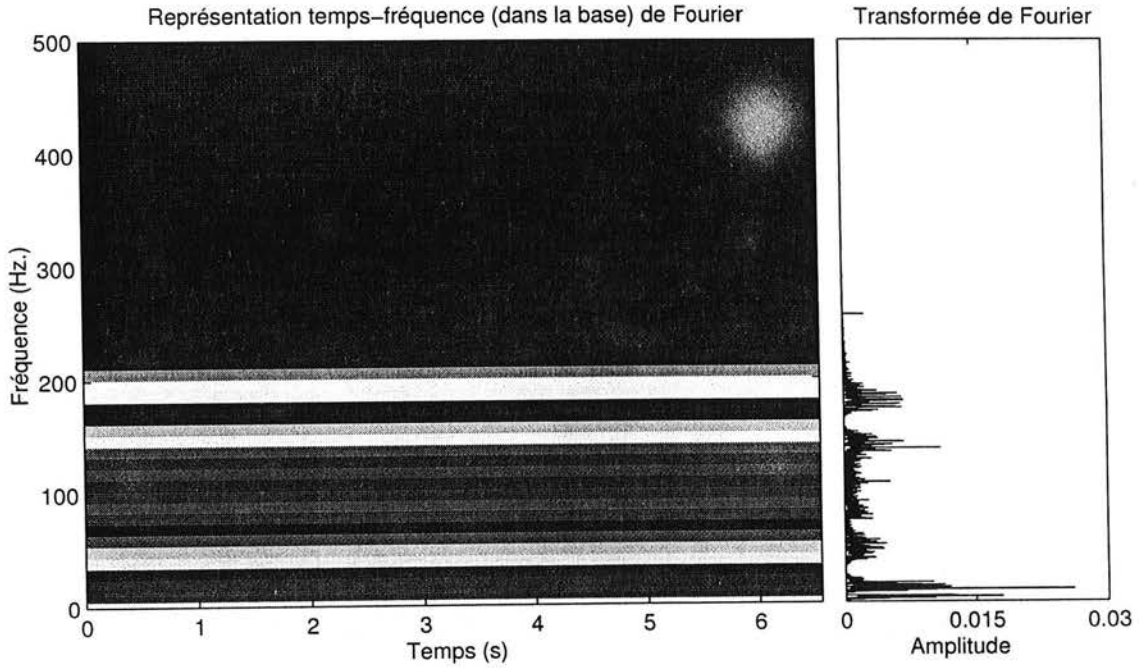


FIG. 3.2 – Représentations fréquentielles (Fourier) 2D et 1D. L'axe fréquentiel est représenté entre 0 et 500 Hz.

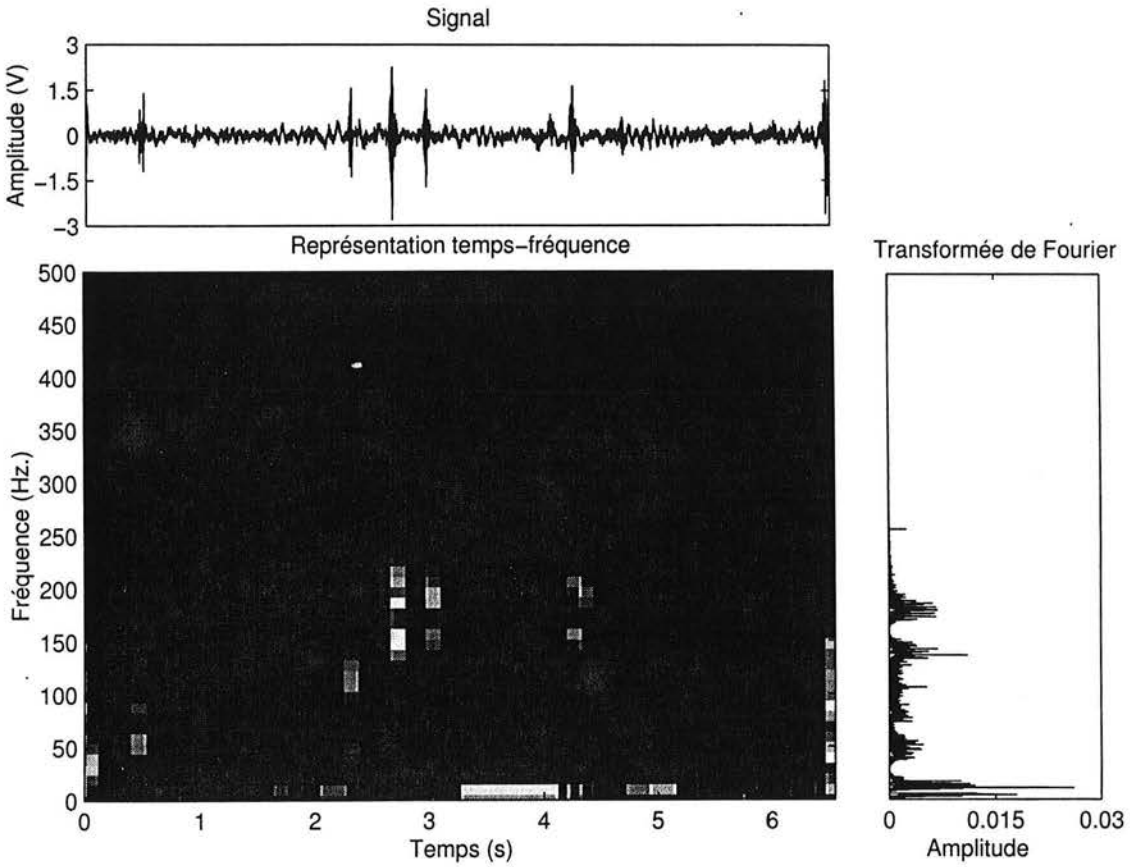


FIG. 3.3 – Représentation temps-fréquence

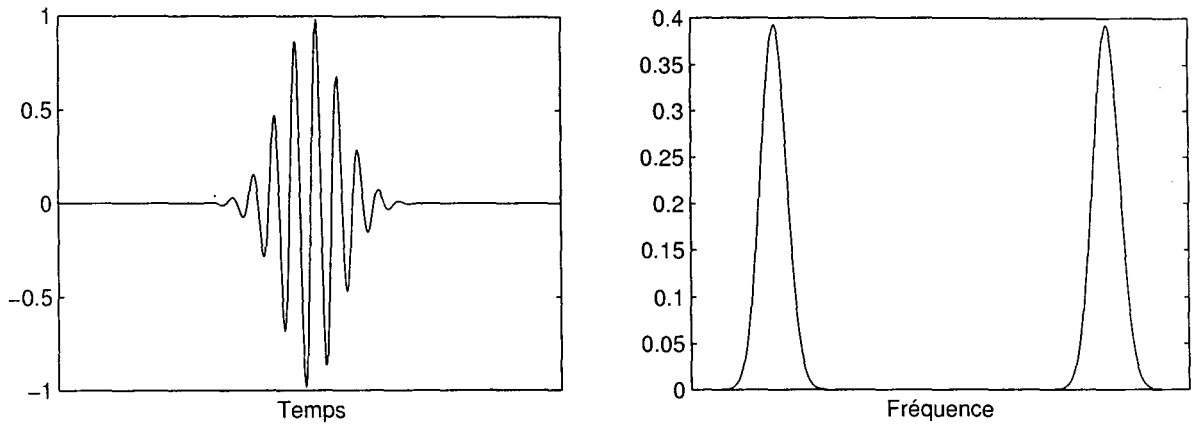
de la limite fixée par le principe d'incertitude[21]. Ces signaux, qui concentrent (d'une manière significative) l'énergie en temps et en fréquence, seront appelés signaux élémentaires, ou *atomes temps-fréquence* (figure 3.4).

Cette section commence avec la présentation d'une des principales représentations temps-fréquence – les décompositions atomiques. Ces décompositions permettent une représentation sans perte du signal qui est en même temps peu ou pas redondante. L'autre grande classe de représentations temps-fréquence, les distributions d'énergie, ne sera pas présentée ici, surtout à cause de son caractère redondant, qui conduit à des charges de calcul très importantes. On estime aussi que sa nature énergétique (donc quadratique) impose des traitements supplémentaires [8, 43], et que ces inconvénients l'emportent sur la précision qu'elle assure, d'autant plus qu'elle n'est pas indispensable à notre application.

3.1.1 Décompositions atomiques

L'idée fondamentale des représentations atomiques est qu'on peut écrire un signal $x(t)$ comme une somme de signaux élémentaires, les atomes temps-fréquence $a_{\tau,\xi}(t)$, localisés dans une fenêtre temporelle de largeur T autour de l'instant t et dans une bande fréquentielle de largeur B autour

[21] Cette limite est atteinte pour une seule famille de signaux, les atomes temps-fréquence de Gabor, c'est-à-dire les exponentielles complexes modulées en amplitude par une gaussienne. Cette approche est équivalente à la transformée de Fourier à court terme.



(a) atome temps-fréquence

(b) spectre correspondant

FIG. 3.4 – Exemple d'atome temps-fréquence de Gabor

d'une fréquence ν :

$$x(t) = \sum_{\tau,\xi} X_{\tau,\xi} \cdot a_{\tau,\xi}(t), \tag{3.1}$$

où les $X_{\tau,\xi}$ sont des coefficients de pondération[22]. Dans le cas général, τ et ξ sont des variables continues. Autrement dit, les pavés temps-fréquence seront superposés, la somme de l'équation (3.1) deviendra une intégrale continue en τ et en ξ et la représentation sera redondante.

Un des intérêts principaux des décompositions atomiques est la possibilité qu'elles offrent de construire des bases, notamment des bases orthonormées. L'équation (3.1) peut être vue comme la représentation d'un vecteur (signal) par la somme de ses projections sur les vecteurs de base (les atomes temps-fréquence), les coefficients $X_{\tau,\xi}$ étant les valeurs scalaires de ces projections. Une propriété intéressante des bases est la complétude, qui assure que la décomposition est d'une taille minimale et donc adaptée au traitement des signaux de grande taille : les pavés du plan temps-fréquence seront disjoints, mais la couverture de ce plan complète [57, 124]. De plus, si la base est orthonormée, elle assure la conservation de l'énergie par décomposition :

$$\sum_{t} |x(t)|^2 = \sum_{\tau,\xi} |X_{\tau,\xi}|^2, \tag{3.2}$$

ce qui se traduit, pour la base constituée des exponentielles complexes définies sur $\mathbb{R}$ (donc pour la transformation de Fourier) par le théorème de Parseval.

En conséquence, pour construire une base d'atomes temps-fréquence, on cherche à éliminer la redondance et à assurer l'orthogonalité. La solution la plus utilisée est le recours aux ondelettes orthogonales[23].

[22] Vu les limitations imposées par le principe d'incertitude de Heisenberg-Gabor, nous utilisons ξ au lieu de ν pour désigner les fréquences localisées dans la bande B. Le même raisonnement est fait pour remplacer t par τ.

[23] D'autres solutions ont été décrites [26, 53, 71, 124] : les bases trigonométriques locales, les paquets d'ondelettes, les ondelettes bi-orthogonales, les représentations redondantes de type *structures obliques* ou *frames*.

3.1.2 Bases d'ondelettes orthogonales

La transformée de Fourier à court terme, les atomes temps-fréquence de Gabor et les bases trigonométriques locales partagent la même approche : à partir d'une série de fonctions orthogonales définies sur $\mathbb{R}$, on multiplie les fonctions de cette série par une fenêtre à support compact (ou rapidement décroissante à $-\infty$ et ∞) glissante en temps. On crée ainsi les atomes temps-fréquence sur lesquels on décompose le signal.

Ce n'est pas le seul moyen pour construire des bases ou des structures obliques. On peut imaginer des fonctions qui sont bien localisées en temps et en fréquence sans être obtenues par la multiplication d'une fonction précise en fréquence (l'exponentielle complexe ou le cosinus) par une fonction fenêtre précise en temps (gaussienne, fenêtre sinusoïdale, Hamming, Kaiser etc.). C'est notamment le cas des ondelettes.

Il faut remarquer que la localisation fréquentielle des ondelettes est une caractéristique secondaire. En effet, on caractérise une décomposition en ondelettes comme une *analyse temps-échelle* et non pas comme une analyse temps-fréquence. Une ondelette est une onde qui oscille un nombre fixé de fois, sans rapport avec la durée. Intuitivement, l'échelle d'une ondelette est proportionnelle à sa durée et inversement proportionnelle à sa "fréquence". On a fait figurer la "fréquence" entre guillemets parce que l'ondelette n'est pas, en général, périodique (voir, par exemple, l'ondelette de Daubechies 2 ou l'ondelette de Haar), donc la fréquence telle qu'elle est utilisée dans la physique et dans l'analyse de Fourier n'a pas de correspondance pour les ondelettes. On va renoncer ensuite aux guillemets, en sous-entendant toujours cette différence entre les ondelettes et les bases de type Fourier.

On appelle ondelette mère une fonction $\psi(t)$ dans $\mathbf{L^2}$ qui a un certain nombre de caractéristiques :

- moyenne nulle :

$$\int_{-\infty}^{\infty} \psi(t)dt = 0 \; ; \tag{3.3}$$

- norme (L_2) unitaire :

$$||\psi(t)||_2 = 1 \; ; \tag{3.4}$$

- centrée autour de $t = 0$.

Par dilatations et translations, on obtient une famille d'ondelettes :

$$\psi_{s,u} = \frac{1}{\sqrt{s}}\psi\left(\frac{t-u}{s}\right), \tag{3.5}$$

où s est l'échelle et u la position temporelle. Si la fonction ψ est régulière (elle est dérivable p fois), sa transformée de Fourier tend vers 0 quand ω tend vers ∞. Comme pour $\omega = 0$ la transformée de Fourier est nulle (directement de (3.3)), l'ondelette est associée à un filtre passe-bande et permet une localisation en temps et en fréquence (la limite est fixée par le principe d'incertitude de Heisenberg). Une décomposition en ondelettes discrète est obtenue en considérant des dilatations s et des translations u discrètes.

Si on considère une suite d'échelles s_j finie, $s_j \leq s_M$, on associe à la famille d'ondelettes (3.5) une famille de fonctions d'échelle, obtenue par translation d'une fonction $\phi(t)$ de norme L_2 unitaire, qui est l'agrégation des ondelettes d'échelle supérieure à s_M [71]. Si les ondelettes sont associées aux filtres passe-bande correspondant aux échelles, la fonction d'échelle est associée à un filtre passe-bas.

La décomposition discrète la plus courante est la décomposition dyadique [71], pour laquelle le pas de discrétisation fréquentiel est l'octave ($s = 2^j$, avec $j = 0 \ldots M$). Le pas temporel p est

lié à l'échelle et dépend de la fonction de base, c'est-à-dire l'ondelette choisie. On notera par la suite j l'échelle et p et la position temporelle.

Une décomposition dyadique d'un signal (une fonction) $x(t)$ sera une somme pondérée de ces ondelettes et fonctions d'échelle. Les coefficients de pondération sont donnés par la transformée en ondelettes, qui est la projection de x sur les ondelettes et les fonctions d'échelle :

$$\begin{aligned} w^{j,p} &= \langle x, \psi_{j,p} \rangle &&= \int_{-\infty}^{\infty} x(t) \frac{1}{\sqrt{2^j}} \psi^* \left(\frac{t-p}{2^j} \right) dt, \\ w^{M,p} &= \langle x, \phi_{M,p} \rangle &&= \int_{-\infty}^{\infty} x(t) \frac{1}{\sqrt{2^M}} \phi^* \left(\frac{t-p}{2^M} \right) dt. \end{aligned} \tag{3.6}$$

Si les fonctions ψ et ϕ (les ondelettes et les fonctions d'échelle) sont orthogonales, on obtient une décomposition orthogonale (une analyse multi-résolution), constituée d'un espace d'approximation et plusieurs espaces de détails. L'espace d'approximation est représenté par la fonction d'échelle (basse fréquence). Inversement, les espaces de détails, caractérisés par les ondelettes d'échelle inférieure, représentent les événements de fréquence supérieure et plus courts. Les espaces de détails sont orthogonaux entre eux et sur l'espace d'approximation.

Afin d'obtenir une transformation rapide et non-redondante du signal, l'algorithme le plus souvent utilisé est la décomposition discrète dyadique sur une base orthonormée d'ondelettes de Mallat [71] (voir figure 3.5), qui assure une décomposition simple et une reconstruction exacte.

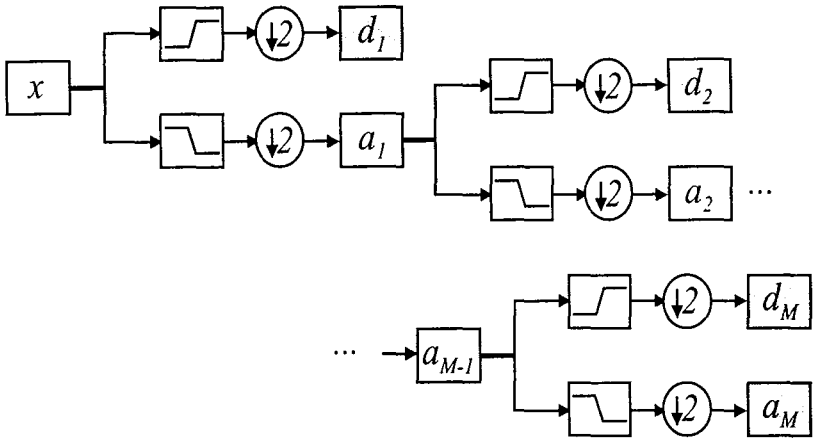


FIG. 3.5 – Algorithme de décomposition en ondelettes de Mallat [71]. Le vecteur des coefficients d'ondelette est obtenu par $\mathbf{w} = \{w^{M,p} \; w^{j,p}\} = \{a_M \; d_M \; d_{M-1} \; \ldots \; d_2 \; d_1\}$.

La base de projection est constituée d'ondelettes $\psi_{j,p}$ et de fonctions d'échelle $\phi_{M,p}$, où j désigne l'échelle (liée à la fréquence), p l'indice (lié à la position temporelle)[24] et M la profondeur de la décomposition en ondelettes. Les coefficients de la décomposition d'un signal x sur cette base sont notés $w^{j,p}$, et le signal s'écrit comme :

$$x(t) = \sum_{j,p} w^{j,p} \psi_{j,p}(t) + \sum_{p} w^{M,p} \phi_{M,p}(t). \tag{3.7}$$

[24]L'indice p par lequel nous notons la position du coefficient d'ondelette est spécifique à l'échelle. Ainsi, pour chaque échelle j, il varie de 1 à N^j, où N^j est le nombre de coefficients de l'échelle considérée.

3.2 Détection, estimation et débruitage par ondelettes

Une des applications les plus courantes des ondelettes est l'extraction du signal informatif c du signal mesuré x. Les points de vue de la communauté scientifique sur cette extraction varient en fonction du domaine et de l'application : on parle de détection quand le but est de localiser les transitoires, d'estimation quand le but est de trouver un modèle du signal informatif, et de débruitage quand on souhaite éliminer le bruit additif r. Cependant, les trois approches se révèlent être des vues différentes d'un même problème : la sélection des coefficients d'ondelette du signal informatif parmi l'ensemble des coefficients du signal mesuré. Par la suite, nous allons parler de débruitage, sous-entendant l'équivalence des points de vue décrits antérieurement.

Dans nos travaux, nous avons considéré un modèle discretisé classique ([6, 38, 77, 83]) du signal mesuré $\boldsymbol{x}$: un mélange additif d'un signal informatif $\boldsymbol{c}$ et d'un bruit de mesure $\boldsymbol{r}$:

$$x(k) = c(k) + r(k). \tag{3.8}$$

La première étape de traitement est la décomposition de ce signal sur une base orthogonale d'ondelettes. Le choix de la base de projection est généralement fait par l'utilisateur[25], mais peut être aussi le résultat d'un algorithme de recherche de la meilleure base, comme proposé par Coifman et Wickerhauser dans [27, 28, 29].

D'après la définition générale de Coifman et Wickerhauser [28, 29], le débruitage par ondelettes revient à "l'extraction d'une structure cohérente" $\boldsymbol{c}$ du signal traité, ce qui, inversement, revient à considérer le bruit $\boldsymbol{r}$ comme non cohérent par rapport à la base d'ondelettes choisie, donc non-corrélé avec les fonctions de base. En fait, les segments du signal peu corrélés avec la base ont des coefficients $w^{j,p}$ faibles, et sont attribués au bruit. Par un seuil adapté, on peut séparer le bruit (partie incohérente) du signal (partie cohérente). Le débruitage par ondelettes classique est donc implémenté comme un filtrage non-linéaire (une régression non-linéaire, selon Antoniadis *et al.* [6]) par seuillage : les coefficients d'ondelette supérieurs à un seuil S sont considérés comme faisant partie du signal informatif.

Nous allons présenter dans la suite les principales techniques de seuillage et d'estimation de la valeur du seuil.

3.2.1 Stratégies de seuillage

Dans la littérature, on trouve plusieurs méthodes de seuillage, dont les plus connues et appliquées sont le seuillage dur (*hard thresholding*) et le seuillage doux (*soft thresholding*) [6, 37, 38, 71, 77]. Le premier propose l'annulation de toutes les valeurs en dessous d'un seuil S, les valeurs supérieures restant inchangées :

$$w_c^{j,p} = \begin{cases} w^{j,p}, & |w^{j,p}| \geq S, \\ 0, & |w^{j,p}| < S. \end{cases}$$

La seconde méthode opère, en plus de l'annulation, une soustraction du seuil S aux valeurs restantes, supérieures au seuil, afin de réduire le nombre de discontinuités dans le signal débruité :

$$w_c^{j,p} = \begin{cases} sign(w^{j,p}) \cdot (|w^{j,p}| - S), & |w^{j,p}| \geq S, \\ 0, & |w^{j,p}| < S. \end{cases}$$

[25] En pratique, pour les sons abdominaux, Hadjileontiadis *et al.* [54, 55] réalisent une transformation en ondelettes sur une base orthonormée de Daubechies4. Le choix de cette base est validé par une expertise médicale auditive après débruitage et reconstruction.

Selon le type de bruit, blanc ou coloré, il est possible d'adapter la stratégie de seuillage. En effet, ce seuillage peut être identique pour tous les coefficients d'ondelette ou, puisque les coefficients sont obtenus par filtrage passe-bande [71] (voir figure 3.5), adapté à chaque échelle. D'autres techniques de seuillage, moins utilisées, sont locales en temps, ou traitent les coefficients par blocs et non pas individuellement (en prenant en compte par exemple l'inter-corrélation entre les échelles adjacentes) [6, 77, 128]. Quelle que soit la stratégie de seuillage, le signal débruité $\mathbf{c}$, *a priori* sans composantes incohérentes, est obtenu par reconstruction à partir des coefficients d'ondelettes supérieurs au seuil (global ou adapté).

3.2.2 Calcul du seuil

Comme dans toute approche de type seuillage, le problème principal réside dans le calcul du seuil S. Un passage en revue très complet des méthodes de calcul est présenté par Antoniadis *et al.* dans [6].

Globalement, on peut distinguer deux grandes classes de méthodes : bayesiennes et classiques. Les méthodes bayesiennes font des hypothèses *a priori* sur les distributions des coefficients d'ondelette du signal informatif et du bruit : la distribution des coefficients du signal mesuré est modélisée comme un mélange de deux distributions de probabilité, et le seuil vise à les séparer. Ces méthodes sont peu appliquées en pratique, leur complexité de calcul étant prohibitive pour des signaux à grand nombre d'échantillons [6].

Les méthodes classiques calculent le seuil (ou les seuils, pour les approches locales en temps ou en fréquence) à partir de la distribution empirique des coefficients d'ondelette, en faisant éventuellement des hypothèses sur le bruit. De plus, implicitement ou explicitement, les coefficients dus au signal sont considérés rares (*sparse*), et certains auteurs proposent des évaluations quantitatives de leur densité par rapport à ceux du bruit : Polchlopek et Noonan [83] introduisent une mesure de rareté (*sparsity*) basée sur la probabilité *a priori* d'apparition d'un grand coefficient, dû au signal informatif, alors que Donoho et Johnstone [37, 38] proposent une méthode heuristique qu'ils utilisent ensuite pour choisir entre deux types de seuils, *universel* et *SURE*.

Nous présentons ensuite les méthodes classiques de débruitage, afin de déterminer la plus adaptée aux sons abdominaux. Ces méthodes peuvent à leur tour être partagées en deux classes, selon la façon de calculer le seuil : directes et itératives.

Méthodes directes

Les méthodes directes se déroulent en général en deux étapes :

1. Calcul du seuil (ou des seuils, échelle par échelle) à partir de la distribution des coefficients d'ondelette.
2. Seuillage et reconstruction du signal débruité à partir des coefficients significatifs.

En fait, les méthodes les plus connues de débruitage par ondelettes sont celles introduites par Donoho et Johnstone [37, 38] : *VisuShrink* et *SureShrink*. La première propose un seuil *universel* S_u déterminé à partir de l'énergie estimée du bruit. Cette méthode utilise l'hypothèse d'un bruit blanc gaussien superposé au signal. Le seuil S_u est égal à $\sigma\sqrt{2\log_e N}$, où N désigne le nombre de points du signal et σ^2 la variance du bruit. Puisque le bruit est considéré blanc, son énergie est équi-répartie sur toutes les bandes de fréquence (les échelles) de la décomposition. En conséquence, l'écart-type est estimé dans une bande où le signal informatif est considéré inexistant ou quasi-inexistant, notamment dans les plus hautes fréquences. Les auteurs utilisent un estimateur robuste de σ à partir de la valeur médiane $\theta_{w^{1,p}}$ des coefficients d'ondelette de la

première échelle :

$$\hat{\sigma} = \frac{\theta_{w^{1,p}}}{0,6745}. \tag{3.9}$$

Cette estimation est appliquée, si la distribution du bruit est gaussienne, pour éviter le biais dû aux grandes valeurs (les coefficients du signal informatif) de l'estimateur classique :

$$\sigma = \sqrt{\frac{1}{N^1}\sum_{p=1}^{N^1}(w^{1,p})^2}, \tag{3.10}$$

où $N^1 = N/2$ est le nombre de coefficients d'ondelettes de l'échelle haute-fréquence.

Dans [11, 71], on montre que ce seuillage est asymptotiquement optimal : la probabilité d'avoir des coefficients supérieurs à $S = \sigma\sqrt{2\log_e N}$ tend vers 0 quand N tend vers ∞. Ce seuil est performant si le signal informatif est creux, c'est-à-dire si ses coefficients sont rares (*sparse*). Autrement dit, le seuillage risque d'être trop fort pour un signal quelconque[26]. Un autre problème posé par le seuillage universel est la valeur unique du seuil pour toutes les bandes de fréquence de la décomposition en ondelettes ou, autrement dit, sa difficulté à prendre en compte un bruit coloré.

Pour remédier à ces problèmes, Donoho et Johnstone ont proposé l'algorithme *SureShrink*, qui adapte les seuils à chaque échelle et les diminue en même temps. Ce seuillage (*SURE - Stein Unbiased Risk Estimator*) est utilisé dans certaines applications pour réduire le risque de non-détection, mais dans notre cas ce n'est pas le problème des non-détections qui se pose, mais plutôt celui des fausses-alarmes. En effet, comme nous l'avons précisé antérieurement (2.1), les signaux sonores abdominaux peuvent être caractérisés comme creux (*sparse*), avec très peu de coefficients de valeurs importantes. Un seuillage haut est donc plus adapté pour les détecter et minimiser en même temps le nombre de fausses-alarmes.

Dans leur article de synthèse, Antoniadis *et al.* [6] présentent aussi d'autres améliorations des méthodes de calcul du seuil décrites dans la littérature : l'invariance par rapport aux translations du signal [25], la validation croisée [80] etc., qui ne seront pas détaillées ici.

Méthodes itératives

Les méthodes itératives de calcul du seuil ont été introduites indépendamment par plusieurs auteurs, dans différents domaines d'application. Le déroulement général de ces méthodes est le suivant (voir figure 3.6) :

1. Initialisation : quelles que soient l'échelle j et la position p, les coefficients du signal informatif sont considérés nuls $\{w_{c,0}^{j,p}\} = 0$ et les coefficients du bruit sont initialisés par les coefficients du signal mesuré $\{w_{r,0}^{j,p}\} = \{w_x^{j,p}\}$.
2. Estimation du seuil : pour l'itération l, le seuil est calculé par différentes méthodes à partir de l'estimation courante du bruit $\boldsymbol{r}_l$ ou de ses coefficients d'ondelette $\{w_{r,l}^{j,p}\}$.
3. Seuillage : les coefficients significatifs (supérieurs au seuil) de l'estimation courante du bruit $\{w_{r,l}^{j,p}\}$ sont attribués au signal informatif, alors que les autres sont attribués à la nouvelle estimation du bruit $\boldsymbol{r}_{l+1}$.
4. Si un critère donné d'arrêt n'est pas atteint, retour au pas 2.

Un des premiers algorithmes, à notre connaissance, est celui de Starck *et al.* [12, 107, 108], développé pour le traitement des images astronomiques. Les auteurs proposent un calcul du seuil

[26] Dans [6], les auteurs énumèrent des seuils encore plus élevés : $\sigma\sqrt{4\log_e N}$ ou $\sigma\sqrt{6\log_e N}$.

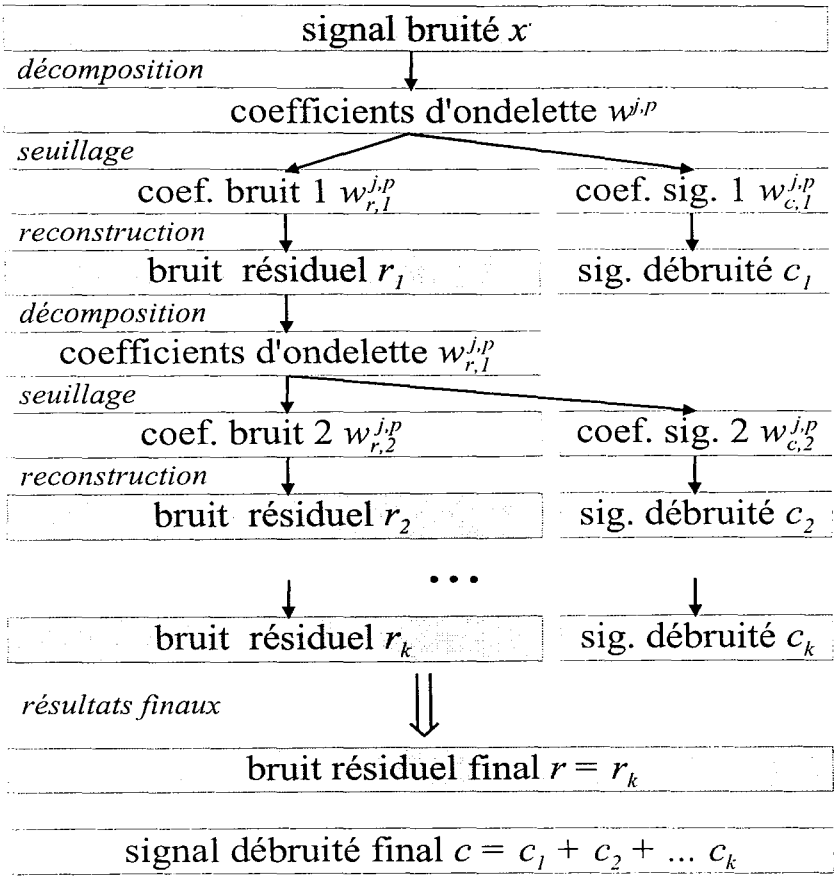


FIG. 3.6 – Schéma de principe d'un algorithme itératif de débruitage, selon Coifman et Wickerhauser [28, 29]. A chaque itération, on annule les coefficients éliminés par seuillage, avant de reconstruire le bruit résiduel et le signal débruité.

(pas 2 de l'algorithme décrit plus haut) basé sur l'écart-type estimé des coefficients d'ondelette du bruit $\{w^{j,p}_{r,l}\}$, échelle par échelle : $S^j_l = 3\sigma^j_l$. L'hypothèse primordiale est la nature gaussienne et blanche du bruit : ainsi, l'écart-type est estimé, comme dans les travaux de Donoho et Johnstone [37, 38], sur la première échelle de la décomposition (la plus haute fréquence), et les coefficients significatifs sont ceux supérieurs en valeur absolue à trois fois l'écart-type. A chaque itération, les "structures suffisamment grandes dans le bruit résiduel" r_l [108] sont extraites et ajoutées à l'estimation du signal informatif, et l'algorithme s'arrête quand aucune information supplémentaire n'est extraite du bruit.

Une approche différente a été introduite par Coifman et Wickerhauser [28, 29]. Les auteurs ne préconisent pas un seuillage en fonction d'un paramètre statistique estimé à partir du signal, ni une hypothèse *a priori* sur la nature du bruit. A la place, ils proposent deux paramètres utilisateur : un seuil $\delta \in [0,1]$ qui sélectionne une proportion $\delta \cdot N$ du nombre total N de coefficients d'ondelette, et un critère d'arrêt β, appelé "coût informationnel" (*information cost*), qui est le nombre maximal autorisé de coefficients du signal informatif. Ce coût informationnel est évidemment lié à la notion *a priori* de densité du signal (*sparsity*). Cependant, l'apport principal de la méthode de Coifman et Wickerhauser est l'intégration de l'algorithme de recherche de la meilleure base de décomposition dans le débruitage, qui permet une efficacité maximale et une compression optimale du signal. Cette amélioration se traduit par contre par une augmentation importante de la charge de calcul.

Enfin, dans le domaine qui nous intéresse, le débruitage itératif par ondelettes des signaux sonores physiologiques pulmonaires et abdominaux a été introduit par Hadjileontiadis *et al.* [54, 55]. L'algorithme WTST-NST (*Wavelet Transform-based Stationary-Nonstationary Filter*) proposé est très proche de celui de Starck *et al.* , mais les auteurs ne font pas d'hypothèse sur la normalité du bruit. Néanmoins, le seuil proposé est identique, la justification de Hadjileontiadis *et al.* étant l'expertise médicale.

A notre connaissance, l'algorithme WTST-NST de débruitage itératif après décomposition sur une base orthonormée unique proposé par Hadjileontiadis *et al.* est le seul expertisé et validé dans le domaine de la phonoentérographie. Vu la nature des signaux sonores abdominaux (décrite dans la section 2.1), notamment leur longue durée, leur caractère creux, avec des événements rares, sous forme de transitoires impulsionnels de contenu fréquentiel relativement pauvre, nous considérons que cet algorithme est en principe adapté à nos besoins.

Cependant, malgré l'utilisation d'un algorithme rapide de décomposition en ondelettes et d'une base de projection unique, la charge de calcul reste importante, surtout pour les signaux de longue durée. De plus, à l'exception notable de la validation médicale, aucune autre méthode de choix des paramètres n'est indiquée à l'utilisateur, ce qui limite l'application de l'algorithme à des signaux acquis en conditions similaires ou proches de celles des auteurs. Notre travail reprend cet algorithme, étudie les aspects théoriques et propose des méthodes d'optimisation et de paramétrage. L'algorithme général que nous proposons est ensuite validé sur des signaux simulés et ses performances sont comparées à d'autres algorithmes de débruitage par ondelettes. Enfin, il est appliqué pour le débruitage et la segmentation des signaux sonores abdominaux réels.

3.3 Algorithme de débruitage de Hadjileontiadis

Dans cette partie nous présentons et réinterprétons la méthode de Hadjileontiadis *et al.* [54, 55] en exploitant les propriétés de la transformée en ondelettes.

Considérons le modèle de l'équation (3.8), $\boldsymbol{x} = \boldsymbol{c} + \boldsymbol{r}$, où $\boldsymbol{x}, \boldsymbol{c}$ et $\boldsymbol{r}$ sont des vecteurs de dimension N. Les méthodes itératives de débruitage réécrivent cette équation comme $\boldsymbol{x} = \boldsymbol{c}_l + \boldsymbol{r}_l$, où l est l'indice d'itération. L'estimation courante du bruit $\boldsymbol{r}_l$ (initialisée à $\boldsymbol{r}_0 = \boldsymbol{x}$) est décomposée sur une base d'ondelettes orthogonale :

$$\boldsymbol{r}_l = \sum_{j,p} w_{r,l}^{j,p} \psi^{j,p} + \sum_{p} w_{r,l}^{M,p} \phi^{M,p}, \tag{3.11}$$

où j est l'échelle, p l'indice de position temporelle et M la profondeur de la décomposition.

Notons $\boldsymbol{\Omega}_{r,l}^{j}$ le vecteur des coefficients du bruit $\{w_{r,l}^{j,p}\}$ à l'échelle j et $\boldsymbol{\Omega}_{r,l}$ le vecteur complet des coefficients du bruit[27]. Par seuillage de $\boldsymbol{\Omega}_{r,l}$, on obtient $\boldsymbol{\Omega}_{\Delta c,l+1}$, le vecteur des grands coefficients qui sont considérés comme appartenant au signal informatif (voir figure 3.7). Le vecteur des coefficients du bruit $\boldsymbol{\Omega}_{r,l+1}$, obtenu à partir de l'équation $\boldsymbol{\Omega}_{\Delta c,l+1} + \boldsymbol{\Omega}_{r,l+1} = \boldsymbol{\Omega}_{r,l}$, est utilisé pour la reconstruction de l'estimation courante du bruit résiduel $\boldsymbol{r}_{l+1}$, qui sert à calculer le critère d'arrêt :

$$STC_{l+1} = \|\boldsymbol{r}_l\|^2 - \|\boldsymbol{r}_{l+1}\|^2 < \varepsilon, \tag{3.12}$$

où ε est un premier paramètre choisi par l'utilisateur.

[27] Vu que le signal est de longueur finie N, les vecteurs de coefficients d'ondelettes sont aussi de dimension finie. En gérant les effets de bord de la décomposition par la périodisation du signal, la dimension des vecteurs de coefficients d'ondelettes sera N, égale à celle du signal.

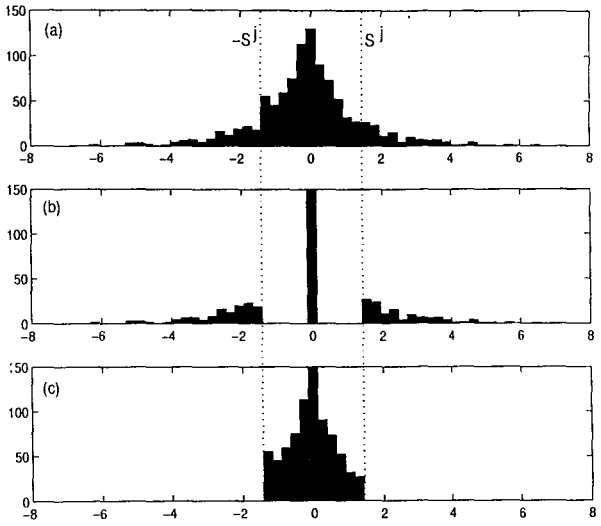


FIG. 3.7 – Histogramme des coefficients d'ondelette à l'échelle j (a) de l'estimation courante du bruit résiduel à l'itération l $\mathbf{\Omega}^j_{r,l}$, (b) du signal informatif $\mathbf{\Omega}^j_{\Delta c,l+1}$, (c) de la nouvelle estimation du bruit résiduel $\mathbf{\Omega}^j_{r,l+1}$. Pour cet exemple, les coefficients d'ondelette du signal simulé suivent une loi de Laplace (l'axe des ordonnées n'est pas complètement représenté en (b) et (c)).

Le vecteur des coefficients informatifs $\mathbf{\Omega}_{\Delta c,l+1}$ contribue à la reconstruction de $\boldsymbol{c}_{l+1}$, qui sert à calculer le signal informatif final (voir figure 3.6).

Un deuxième paramètre utilisateur F_a intervient dans le calcul du seuil S. L'algorithme WTST-NST de Hadjileontiadis *et al.* [54, 55] calcule ce seuil en fonction de l'écart-type σ^j_l des coefficients d'ondelette $\{w^{j,p}_{r,l}\}$ du bruit estimé à chaque échelle. Le seuil est recalculé, indépendamment pour chaque échelle, à chaque itération l :

$$S^j_{l+1} = F_a \cdot \sigma^j_l, \tag{3.13}$$

où l'expression de la constante $F_a = 3$ est validée par expertise médicale[28].

Schématiquement, l'algorithme de Hadjileontiadis *et al.* [54, 55] s'écrit :

1. calculer σ^j_l comme $\left(\sigma^j_l\right)^2 = \frac{1}{N^j}\left\|\mathbf{\Omega}^j_{n,l}\right\|^2$, où $N^j = card\left(\mathbf{\Omega}^j_{n,l}\right)$ est le nombre de coefficients de l'échelle j ($N^j = \frac{N}{2^j}$ pour $j = 1 \ldots M$ et $N^j = \frac{N}{2^M}$ pour $j = M+1$) ;
2. calculer le seuil S^j_{l+1} comme $S^j_{l+1} = f(\sigma^j_l)$. Dans l'algorithme présenté ici, $S^j_{l+1} = F_a\sigma^j_l$, où F_a est un paramètre choisi par l'utilisateur et indépendant de l'échelle j ;
3. calculer $\mathbf{\Omega}^j_{r,l+1}$ et $\mathbf{\Omega}^j_{\Delta c,l+1}$ par seuillage dur de $\mathbf{\Omega}^j_{r,l}$ avec le seuil S^j_{l+1} ;
4. calculer $\boldsymbol{r}_{l+1}$ et $\Delta\boldsymbol{c}_{l+1}$ comme les reconstructions par ondelettes des vecteurs $\mathbf{\Omega}_{r,l+1}$ et $\mathbf{\Omega}_{\Delta c,l+1}$, et fixer $\boldsymbol{c}_{l+1} = \boldsymbol{c}_l + \Delta\boldsymbol{c}_{l+1}$;
5. revenir au pas 1 si le critère d'arrêt (3.12) n'est pas atteint.

Cet algorithme itératif s'interrompt quand la différence d'énergie sur le bruit résiduel $\boldsymbol{r}$ reconstruit entre deux itérations successives est jugé suffisamment faible. Un STC_l petit est obtenu quand très peu d'information (parties du signal corrélées avec l'ondelette choisie) est extraite du bruit et ajoutée au signal entre deux itérations successives.

[28] L'expression du seuil est similaire dans les travaux de Starck *et al.* [108] en imagerie astronomique. Cette valeur de la constante F_a est liée aussi, sous hypothèse d'une distribution gaussienne des coefficients d'ondelettes, à la détection des valeurs aberrantes (*outliers*).

3.4 Interprétation de point-fixe

Nous présentons dans cette section une analyse détaillée de l'algorithme WTST-NST de Hadjileontiadis *et al.* . En effet, l'utilisation d'une base de décomposition-reconstruction unique permet des développements intéressants de cet algorithme. Nous montrons notamment que, dans ce cas, les étapes de décomposition-reconstruction successives (voir figure 3.6) sont inutiles, ce qui implique une réduction importante de la charge de calcul. L'étude de la convergence de la méthode itérative de seuillage (selon l'équation 3.13) nous permet d'éliminer un des paramètres utilisateur ε grâce à l'introduction d'une approche de point-fixe [86, 89] et, de ce fait, d'optimiser l'algorithme de débruitage.

3.4.1 Préliminaires

Les transformées en ondelettes[29] directe (WD) et inverse (WR) appliquées pour les opérations de décomposition et reconstruction sont orthogonales et linéaires. En prenant en compte l'orthogonalité de la transformation en ondelettes, le critère d'arrêt (3.12) s'écrit :

$$STC_{l+1} = \|\boldsymbol{r}_l\|^2 - \|\boldsymbol{r}_{l+1}\|^2 = \|\boldsymbol{\Omega}_{r,l}\|^2 - \|\boldsymbol{\Omega}_{r,l+1}\|^2 . \tag{3.14}$$

En choisissant $\varepsilon = 0$, les itérations s'arrêtent quand $\boldsymbol{\Omega}_{r,l} = \boldsymbol{\Omega}_{r,l+1}$, c'est-à-dire quand aucun coefficient n'est éliminé par seuillage et $\boldsymbol{\Omega}_{\Delta c,l+1} = \mathbf{0}$. Le critère d'arrêt (3.14) modifié ($STC_{l+1} = 0$) peut être utilisé à l'étape no. 5 de l'algorithme précédent, ce qui implique que la reconstruction du bruit $\boldsymbol{r}_l = \text{WR}(\boldsymbol{\Omega}_{r,l})$ à chaque itération est inutile.

De plus, puisque la transformation en ondelettes est linéaire, le vecteur des coefficients d'ondelette du signal $\boldsymbol{x}$ s'écrit :

$$\boldsymbol{\Omega}_x = \boldsymbol{\Omega}_{r,l} + \sum_{i=0}^{l} \boldsymbol{\Omega}_{\Delta c,i} = \boldsymbol{\Omega}_{r,l} + \boldsymbol{\Omega}_{c,l}, \ \forall\, l, \tag{3.15}$$

ce qui implique qu'il n'est pas nécessaire de calculer $\Delta \boldsymbol{c}_l$ par reconstruction à partir de $\boldsymbol{\Omega}_{\Delta \boldsymbol{c},l}$ à chaque itération et le signal informatif final $\boldsymbol{c}$ se reconstruit selon $\boldsymbol{c} = \sum_{l=0}^{L} \Delta \boldsymbol{c}_l = \sum_{l=0}^{L} \text{WR}(\boldsymbol{\Omega}_{\Delta c,l})$. Ce signal sera calculé seulement après convergence : $\boldsymbol{c} = \text{WR}(\boldsymbol{\Omega}_c) = \text{WR}(\sum_{l=0}^{L} \boldsymbol{\Omega}_{\Delta c,l})$, où L est le nombre final d'itérations.

Autrement dit, les étapes intermédiaires de décomposition-reconstruction qui constituent l'étape no. 4 de l'algorithme de Hadjileontiadis *et al.* peuvent être supprimées et remplacées par une reconstruction finale unique.

3.4.2 Seuillage

Pour chaque échelle j, le seuillage des coefficients d'ondelette de l'estimation du bruit $\boldsymbol{\Omega}^j_{r,l-1}$ est effectué avec le seuil $S^j_l = F_a \sigma^j_{l-1}$, qui sépare $\boldsymbol{\Omega}^j_{r,l-1}$ en $\boldsymbol{\Omega}^j_{r,l}$ et $\boldsymbol{\Omega}^j_{\Delta c,l}$. S'il existe au moins un indice q tel que $\boldsymbol{\Omega}^j_{\Delta c,l}(q) \neq 0$, alors on a $|\boldsymbol{\Omega}^j_{\Delta c,l}(q)| \geq S^j_l$. En plus, pour tout q, $S^j_l > |\boldsymbol{\Omega}^j_{r,l}(q)|$ et

$$|\boldsymbol{\Omega}^j_{\Delta c,l}(q)| \geq S^j_l > |\boldsymbol{\Omega}^j_{\Delta c,l+1}(q)| \geq S^j_{l+1} > \ldots, \tag{3.16}$$

[29] Tout le long de cette thèse, nous considérons un signal périodisé pour prendre en compte les effets de bord de la transformée. Cette approche permet d'obtenir un nombre de coefficients d'ondelette égal à la taille du signal et détermine par la suite le calcul des indices de ces coefficients.

car $\mathbf{\Omega}^j_{\Delta c,l+1}$ est obtenu par le seuillage de $\mathbf{\Omega}^j_{r,l}$. Par conséquent, tous les coefficients non-nuls de $\mathbf{\Omega}^j_{c,l} = \sum_{i=0}^{l} \mathbf{\Omega}^j_{\Delta c,i}$ ont un module supérieur ou égal à S^j_l, qui est supérieur ou égal aux coefficients de $\mathbf{\Omega}^j_{r,l}$. En initialisant la suite de seuils par $S^j_0 = \infty$ (ce qui est équivalent à l'initialisation $\boldsymbol{r}_0 = \boldsymbol{x}$ proposée précédemment), on obtient (voir (3.16)) une séquence décroissante $S^j_0 > S^j_1 > \cdots > S^j_{L-1} = S^j_L$, où S^j_L est le seuil final obtenu quand $\mathbf{\Omega}_{\Delta c,L}(q) = 0$ quel que soit q.

Par conséquent, S^j_l sépare $\mathbf{\Omega}^j_x = \mathbf{\Omega}^j_{r,l} + \mathbf{\Omega}^j_{c,l}$ en deux vecteurs disjoints quel que soit l et donc le seuillage de $\mathbf{\Omega}^j_{r,l}$ par S^j_{l+1} est équivalent au seuillage direct de $\mathbf{\Omega}^j_x$. Les seuillages successifs de $\mathbf{\Omega}^j_x$ sont inutiles et donc le seuil final S^j_L peut être calculé dans une étape antérieure de l'algorithme.

3.4.3 Interprétation de point fixe

Quel que soit le paramètre de seuillage F^j_a, le seuil S^j_{l+1} (à l'itération $l+1$) est calculé en fonction de l'écart-type des coefficients restants dans le bruit à l'itération précédente, c'est-à-dire des coefficients qui respectent $\left|w^{j,p}_{r,l}\right| < S^j_l$. En choisissant dans l'équation (3.14) $\varepsilon = 0$, l'algorithme s'écrit, pour chaque échelle j :

1. calculer σ^j_l comme $\left(\sigma^j_l\right)^2 = \dfrac{1}{N^j} \sum_p \left(w^{j,p}_z \, \mathbb{I}_{(|w^{j,p}_z| < S^j_l)}\right)^2$, où N^j est le nombre de coefficients à l'échelle j et
$$\mathbb{I}_{(|x|<S)} = \begin{cases} 1, & \text{si } |x| < S, \\ 0, & \text{si } |x| \geq S; \end{cases}$$
2. calculer S^j_{l+1} comme $S^j_{l+1} = F^j_a \, \sigma^j_l$;
3. revenir au pas 1. si la convergence n'est pas atteinte.

On peut donc écrire l'itération suivante :

$$S^j_{l+1} = f_j(S^j_l) = F^j_a \sqrt{\frac{1}{N^j} \sum_p \left(w^{j,p}_x \, \mathbb{I}_{(|w^{j,p}_x| < S^j_l)}\right)^2}, \tag{3.17}$$

$$\text{avec} \quad f_j(u) = F^j_a \sqrt{\frac{1}{N^j} \sum_p \left(w^{j,p}_x \, \mathbb{I}_{(|w^{j,p}_x| < u)}\right)^2}. \tag{3.18}$$

Évidemment, les itérations $S^j_{l+1} = f_j(S^j_l)$ convergent quand $S^j_{l+1} = S^j_l$, ce qui est équivalent à $S^j_l = f_j(S^j_l)$, donc vers un point fixe de la fonction f_j (voir figure 3.8(b)).

En conséquence, on opère un débruitage par descente de point fixe pour chaque échelle indépendamment. On interrompt les itérations quand le critère est atteint pour l'échelle donnée. La reconstruction du bruit résiduel et du signal informatif n'est plus nécessaire à chaque itération.

3.4.4 Conditions générales d'application

On a montré au paragraphe précédent que le critère d'arrêt est validé pour un point fixe de la fonction f_j définie antérieurement (3.18), et que ce point est le seuil final pour l'échelle considérée j. Il reste à étudier dans quelles conditions ce seuil existe, c'est-à-dire les conditions de convergence de l'algorithme ou encore l'existence de ce point fixe de la fonction f_j.

Il est évident que la fonction f_j définie en (3.18) est une fonction monotone croissante constante par morceaux (voir figure 3.8(a)), définie pour tout $u \in \mathbb{R}^+$ et à valeurs dans un

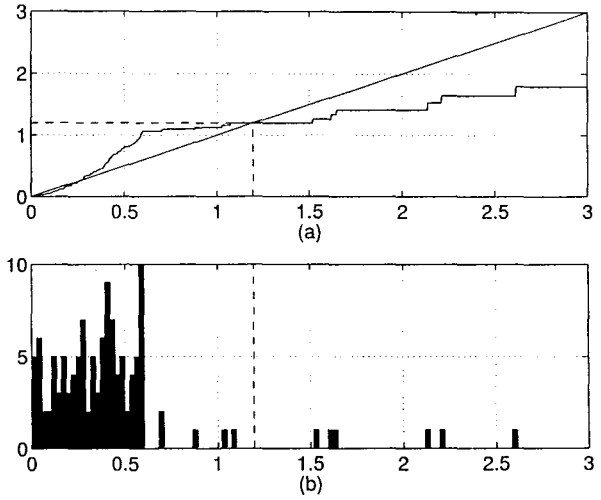


FIG. 3.8 – Pour une distribution empirique de $w_z^{j,p}$ et un j donné : (a) les fonctions $y = f_j(x)$ (pour $F_a^j = 3$) et $y = x$, montrant l'interprétation de point fixe et la valeur finale $x_0 = S_L^j$ donnée par l'algorithme ; (b) l'histogramme des valeurs absolues des coefficients d'ondelette $w_z^{j,p}$. La ligne en pointillé représente le seuil de séparation entre le bruit (les petites valeurs) et le signal informatif.

ensemble fini de nombres réels, étant donné que la dimension du vecteur de coefficients d'ondelette $\mathbf{\Omega}_x^j$ est finie. Les valeurs de la fonction f_j dépendent de la constante utilisateur F_a^j, qui doit respecter certaines conditions pour assurer l'existence du point fixe :

- pour un a donné tel qu'il existe au moins un coefficient $w_x^{j,p}$, $0 < |w_x^{j,p}| < a$, on peut trouver une constante finie F_a^j qui assure $f_j(a) > a$;
- pour le F_a^j trouvé, on peut trouver $b > a$ suffisamment grand pour avoir $f_j(b) < b$.

Par conséquent (voir figure 3.8(a)), on peut affirmer l'existence d'au moins un point fixe $u_f \in \mathbb{R}^+$ tel que $f_j(u_f) = u_f$. Selon l'équation (3.16), l'itération $S_{l+1}^j = f_j(S_l^j)$ génère une suite décroissante de seuils. On peut montrer qu'elle converge vers le premier point fixe rencontré, c'est-à-dire le plus grand :

> Notons $\mathbf{\Gamma}$ l'ensemble fini des valeurs de la fonction f_j (3.18) pour une échelle j donnée, et S_l^j l'ensemble des seuils obtenus par itération ($l \neq 0$). Puisque la suite des seuils est calculée à l'aide de la fonction f_j, $S_l^j \in \mathbf{\Gamma}, \forall\ l \neq 0$. Supposons que le plus grand point fixe de f_j ne soit pas le point de convergence de l'algorithme. Ceci implique qu'il existe $\gamma = f_j(\gamma)$ tel que $\gamma \in \mathbf{\Gamma}$ et $\gamma \neq S_l^j, \forall\ l$ (donc un point fixe non-détecté par l'algorithme itératif). Ce constat peut être réécrit : il existe l tel que $S_{l+1}^j < \gamma < S_l^j$. Mais $S_{l+1}^j = f_j(S_l^j)$, donc la première inégalité s'écrit $f(S_l^j) < \gamma$. D'autre part, f_j est monotone croissante, donc la deuxième inégalité implique $f_j(\gamma) \leq f_j(S_l^j)$, qui implique (car γ est un point fixe) $\gamma \leq f_j(S_l^j)$, ce qui contredit la première inégalité. Par conséquent, l'hypothèse faite au début de notre démonstration est fausse : le plus grand point fixe est le point de convergence de l'algorithme.

Le point fixe trouvé par notre algorithme est donc le seuil final de débruitage pour $\varepsilon = 0$. L'interprétation de l'algorithme de débruitage comme une descente de point fixe permet de mettre en évidence le rôle du deuxième paramètre utilisateur, la constante multiplicative F_a^j :

- elle peut varier d'une échelle à l'autre, en fonction de la distribution des coefficients d'ondelette ;
- elle doit être choisie suffisamment grande pour assurer la convergence de l'algorithme vers une valeur strictement positive (un seuil nul équivaut à considérer qu'il n'y a pas de bruit dans le signal mesuré $\boldsymbol{x}$) ;

- d'autre part, une valeur trop grande de F_a^j conduit à un arrêt prématuré de l'algorithme : si $F_a^j > \frac{\max(|\mathbf{\Omega}_x^j|)}{\sigma_0^j}$, alors l'algorithme converge après la première itération (ce qui signifie qu'il n'y a pas de signal informatif dans la bande fréquentielle correspondant à l'échelle courante j).

3.5 Modélisation probabiliste

Le développement présenté dans la section précédente permet d'éliminer les étapes de décomposition-reconstruction successives et de réinterpréter l'algorithme de débruitage par ondelettes de Hadjileontiadis *et al.* [54, 55] comme un algorithme de point fixe, ce qui conduit à l'élimination des étapes intermédiaires de décomposition-reconstruction et à une nette réduction de la charge de calcul (voir le tableau 3.6, page 77). En plus de l'élimination d'un des paramètres utilisateur (ε), cette approche a mis en évidence l'importance du second paramètre F_a^j. Dans cette section, nous proposons une méthode de calcul de ce deuxième paramètre, ce qui nous permettra de développer une version non-supervisée de l'algorithme de débruitage itératif [89].

Le résultat de la transformée en ondelettes discrète utilisée pour analyser le signal mesuré est un vecteur $\mathbf{\Omega}_x$ de dimension finie, qui peut être divisé en plusieurs vecteurs $\mathbf{\Omega}_x^j$ correspondant à chaque échelle. Les coefficients d'ondelette de chaque échelle peuvent être considérés comme un échantillon discret issu d'une loi de densité de probabilité continue. On observe que, dans la plupart des applications réelles, les coefficients d'ondelette sont distribués selon une loi de probabilité centrée en 0 et qui possède un mode en 0. Notons cette loi $p(w^{j,p})$, ou, puisque le développement suivant est identique pour toutes les échelles, $p(w)$ (pour simplifier la notation, dans la suite on renonce à l'indice j de l'échelle).

On peut alors énoncer la proposition suivante, établie ci-après :

Proposition 1. Si la loi de probabilité des coefficients $p(w)$ de moyenne nulle et de variance finie possède un mode en 0 et s'il y a au moins un point d'intersection entre $p(w)$ et la fonction $h(w) = \frac{K}{w}$, avec $K = \frac{3}{2F_a^2}$, c'est-à-dire que l'équation $p(w) = h(w)$ a au moins une solution réelle, alors la fonction $f(S) = F_a \cdot \sigma_{|w|<S}$ possède au moins un point fixe (voir figure 3.9).

Une fonction $g(x)$ continue sur un intervalle $[a, b]$ possède un point fixe dans $[a, b]$ si $\forall\ x \in [a, b]$, $g(x) \in [a, b]$. Pour une fonction g monotone croissante, $g(x) \in [a, b]$ s'écrit $g(a) > a$ et $g(b) < b$.

Pour une loi de probabilité $p(w)$ paire (ce qui est le cas dans la plupart des applications réelles), la fonction $f(S)$ définie dans (3.18), se redéfinit comme

$$f(S) = F_a \cdot \sigma_{|w|<S} = F_a \sqrt{\int_{-S}^{S} p(w) w^2 dw} = F_a \sqrt{2 \int_{0}^{S} p(w) w^2 dw} \tag{3.19}$$

et est continue quand $p(w)$ est continue.

Afin de vérifier les conditions d'existence d'un point fixe de cette fonction :

1. On montre qu'elle est monotone croissante. Le résultat est immédiat par positivité de $p(w)$.
2. On cherche un intervalle $[S_{fin}, S_0]$ tel que $f(S_0) < S_0$ et $f(S_{fin}) > S_{fin}$. Si le moment d'ordre 2 de $p(w)$ existe, on peut calculer $M = F_a \cdot \sigma_w$. Dans ce cas :

$$\forall\ S_0 > M, f(S_0) = F_a \cdot \sigma_{|w|<S_0} = M < S_0. \tag{3.20}$$

3. Il reste à trouver un point S_{fin} tel que $f(S_{fin}) > S_{fin}$. L'inéquation $f(S_{fin}) > S_{fin}$ s'écrit :

$$F_a^2 \cdot 2 \int_0^{S_{fin}} p(w) w^2 dw > S_{fin}^2. \tag{3.21}$$

En développant le premier membre de (3.21) en intégrant par parties, on obtient :

$$F_a^2 \cdot 2 \int_0^{S_{fin}} p(w) w^2 dw = 2F_a^2 \left[p(w) \frac{w^3}{3} \Big|_0^{S_{fin}} - \int_0^{S_{fin}} p'(w) \frac{w^3}{3} dw \right].$$

Si la fonction de densité de probabilité $p(w)$ a un mode en 0, elle est décroissante sur un intervalle $[0, S_{fin}]$, donc sa dérivée est négative. Dans ce cas, l'intégrale $\int_0^{S_{fin}} p'(w) \frac{w^3}{3} dw$ est négative, donc

$$\begin{aligned} 2F_a^2 \left[p(w) \frac{w^3}{3} \Big|_0^{S_{fin}} - \int_0^{S_{fin}} p'(w) \frac{w^3}{3} dw \right] &= 2F_a^2 \left[p(S_{fin}) \frac{S_{fin}^3}{3} - \int_0^{S_{fin}} p'(w) \frac{w^3}{3} dw \right] \\ &> 2F_a^2 \cdot p(S_{fin}) \frac{S_{fin}^3}{3}. \end{aligned} \tag{3.22}$$

En utilisant cette minoration (3.22), on vérifie que l'inéquation (3.21) est respectée si :

$$2F_a^2 \cdot p(S_{fin}) \frac{S_{fin}^3}{3} \geq S_{fin}^2 \Leftrightarrow 2F_a^2 \cdot p(S_{fin}) \frac{S_{fin}}{3} \geq 1 \Leftrightarrow p(S_{fin}) \geq \frac{3}{2F_a^2 S_{fin}}. \tag{3.23}$$

Si l'inégalité (3.23) est vérifiée, l'inégalité (3.21) est respectée également et on peut affirmer l'existence d'un point S_{fin} tel que

$$f(S_{fin}) > S_{fin}. \tag{3.24}$$

Les relations (3.20) et (3.24) montrent que la fonction monotone croissante $f(S)$ a au moins un point fixe supérieur à 0, si la condition (3.23) est vérifiée. Il est à remarquer que la condition

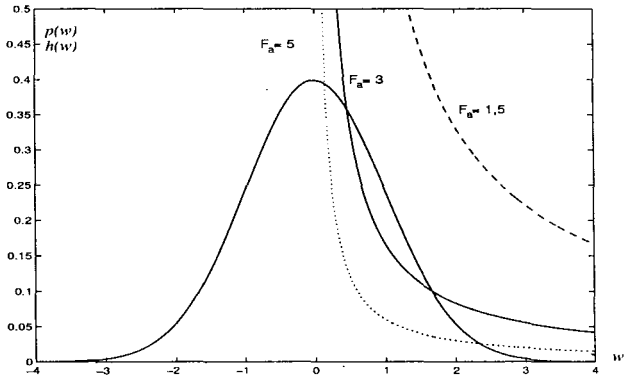


FIG. 3.9 – Condition d'existence du point fixe. Dans cet exemple, la loi de probabilité $p(w)$ est une gaussienne de moyenne nulle et d'écart-type $\sigma = 1$. La fonction $h(w) = \frac{3}{2F_a^2 w}$ est tracée pour trois valeurs de F_a. La convergence de l'algorithme de point fixe n'est pas assurée pour $F_a = 1,5$.

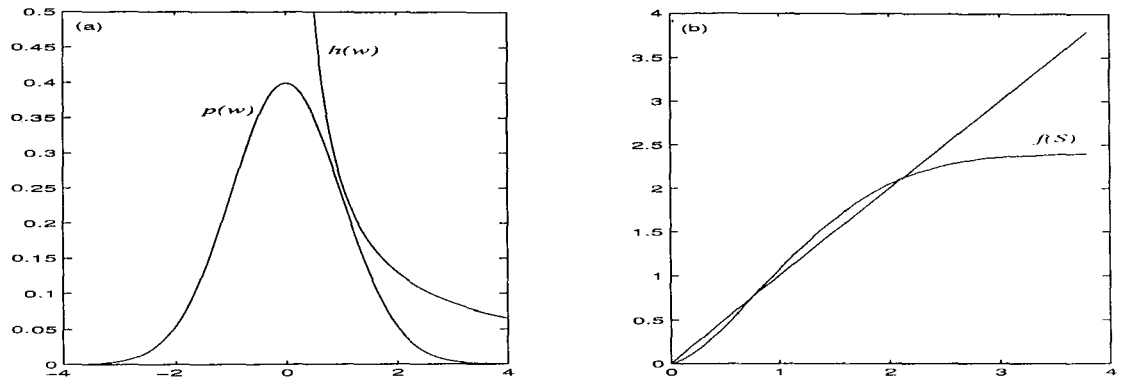


FIG. 3.10 – Condition d'existence du point fixe. (a) l'équation $p(w) = h(w)$ n'a pas de solution ; (b) la fonction $f(S)$ a deux points fixes, donc l'algorithme de débruitage converge. Dans les deux figures, $F_a = 2,4$.

introduite par la proposition précédente est suffisante, mais pas nécessaire, et cela à cause de la minoration (3.22). Par exemple (voir figure 3.10), selon la même loi de probabilité illustrée figure 3.9, mais pour un seuil $S = F_a\sigma$ (avec $F_a = 2,4$), le point fixe de la fonction $f(S)$ existe mais il n'y a aucun point d'intersection entre $h(w)$ et $p(w)$.

Nous avons montré que, sous certaines conditions (qui sont proches des conditions réelles) sur la loi de probabilité des coefficients d'ondelette de l'échelle j, on peut toujours trouver un coefficient F_a^j qui assure la convergence de l'algorithme de débruitage, ce qui implique que le coefficient F_a^j doit être borné par valeur minimale F_{am}^j. En effet, si F_a^j est choisi inférieur à F_{am}^j, l'algorithme converge en 0, ce qui signifie que le bruit estimé est nul à l'échelle considérée.

3.5.1 Modèles gaussiens généralisés

Comme nous l'avons montré précédemment, seule une valeur de F_a supérieure à F_{am} assure la convergence de l'algorithme de débruitage pour l'échelle j considérée. Cette valeur minimale dépend de la distribution de probabilité des coefficients d'ondelette $p(w)$ (3.23)[30]. Nous nous proposons [89] de déterminer l'expression de F_{am} dans le cas des lois de probabilité gaussiennes généralisées, car ce sont des modèles pertinents pour des données expérimentales [112]. Les lois de probabilité de type gaussienne généralisée sont définies selon :

$$p_{\sigma,u}(w) = \alpha e^{-|\beta w|^u}, \tag{3.25}$$

$$\text{avec } \beta = \frac{1}{\sigma}\sqrt{\frac{\Gamma(3/u)}{\Gamma(1/u)}}, \quad \alpha = \frac{\beta u}{2\Gamma(1/u)}, \quad \Gamma(u) = \int_0^\infty e^{-x}x^{u-1}dx,$$

où σ désigne l'écart-type, $u > 0$ le paramètre qui gouverne la décroissance de la loi et sa forme ($u = 2$ pour la loi normale et $u = 1$ pour la loi de Laplace).

On cherche la valeur minimale du paramètre F_a pour laquelle la condition énoncée dans la proposition 1 est respectée. Cela revient à trouver la solution de l'équation $p(w) = h(w)$. L'intersection existe seulement pour w positif et le paramètre β de la gaussienne généralisée est toujours positif, donc l'équation s'écrit :

$$\alpha e^{-(\beta w)^u} = \frac{3}{2F_a^2 w}, \text{ ou encore } w = \frac{3}{2\alpha F_a^2}e^{(\beta w)^u}. \tag{3.26}$$

[30]Pour alléger les notations, on ne fait pas apparaître l'indice j de l'échelle dans le développement présenté ici.

La fonction $q(w) = \frac{3}{2F_a^2} e^{(\beta w)^u}$ est monotone croissante et convexe. On montre aisément qu'elle a 0, 1 ou 2 points d'intersection avec la droite $y(w) = w$, et donc que l'équation (3.26) admet 0, 1 ou 2 solutions. On obtient une solution unique w_0 lorsque $q(w_0)$ est tangente à la première bissectrice. Cela revient à chercher le point w_0 pour lequel on a $q'(w_0) = 1$, c'est-à-dire que :

$$\frac{3}{2\alpha F_a^2} u\beta^u w_0^{u-1} e^{(\beta w_0)^u} = 1. \tag{3.27}$$

En remplaçant (3.26) dans (3.27), on obtient :

$$w_0^u = \frac{1}{u\beta^u}. \tag{3.28}$$

Donc, si on cherche une solution unique de l'équation (3.26), on trouve :

$$w_0 = \frac{1}{\beta u^{\frac{1}{u}}}. \tag{3.29}$$

En remplaçant w_0 en (3.26), on trouve la valeur de F_{am} pour laquelle cette solution existe, et donc la valeur minimale qui assure la convergence de l'algorithme de débruitage de point fixe, en considérant une modélisation de la distribution des coefficients d'ondelette par des gaussiennes généralisées :

$$F_{am} = \sqrt{\frac{3\beta}{2\alpha}(ue)^{\frac{1}{u}}} = \sqrt{\frac{3\Gamma(\frac{1}{u})}{u}(ue)^{\frac{1}{u}}}, \tag{3.30}$$

indépendant de σ mais différent pour chaque échelle j de la décomposition.

3.5.2 Optimisation

Les valeurs minimales F_{am}^j des constantes multiplicatives assurent la convergence de l'algorithme de débruitage de point fixe vers des seuils différents de 0 à chaque échelle j. Cependant, ces seuils ont un désavantage important : les valeurs de F_{am}^j sont souvent inférieures à 3 et elles diminuent avec le paramètre u qui gère la forme des gaussiennes généralisées (figures 3.11 et 3.12). Dans ce cas, le seuil obtenu est inférieur à celui de Hadjileontiadis *et al.* et la version reconstruite du signal risque de contenir trop de fausses alarmes et donc être trop bruitée. D'autre part, un avantage apporté par les seuils calculés à partir de F_{am}^j est qu'ils ne dépendent pas de la taille N du signal, comme dans le cas de l'algorithme *VisuShrink* de Donoho et Johnstone. En effet, le seuil universel $S_u = \sigma\sqrt{2\log_e N}$ tend vers l'infini quand la longueur du signal augmente et donc peu d'événements (voire aucun) seront extraits du bruit pour les signaux de grande taille.

Nous avons donc développé trois versions de notre algorithme de point fixe, qui diffèrent par la méthode de calcul du seuil :

1. **seuillage minimal** : $S_l^j = F_{am}^j \sigma_l^j$, avec F_{am}^j valeur minimale du paramètre F_a^j (éq. (3.30)) qui assure la convergence de l'algorithme de point fixe et σ_l^j écart-type des coefficients d'ondelette du bruit résiduel à l'échelle j et à l'itération l. Ce seuillage conduit en principe à un débruitage sans perte d'information (aucune non-détection, mais certainement de nombreuses fausses alarmes) ;
2. **seuillage "optimal"** : on calcule le paramètre F_{ao}^j tel que les seuils S^j soient initialisés, pour notre algorithme itératif et pour un bruit additif gaussien, aux valeurs correspondant au seuil universel de Donoho et Johnstone :

$$F_{ao}^j = K_o^j F_{am}^j, \tag{3.31}$$

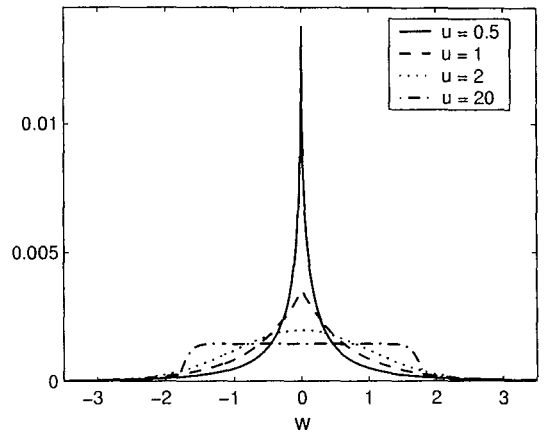


FIG. 3.11 – Exemples de fonctions de densité de probabilité gaussiennes généralisées $p_{\sigma,u}(w)$ de moyenne nulle et d'écart-type $\sigma = 1$, pour différentes valeurs de u.

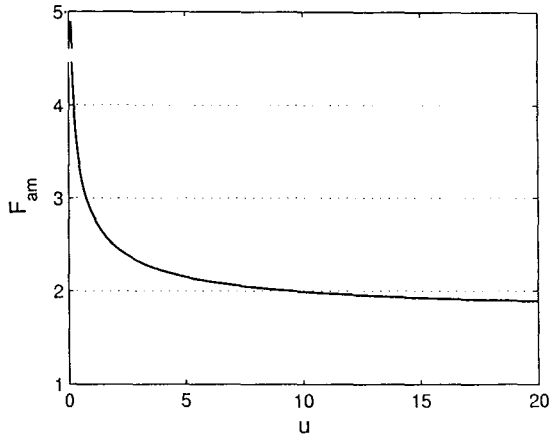


FIG. 3.12 – Variation de la constante multiplicative F^j_{am} en fonction du paramètre u des gaussiennes généralisées.

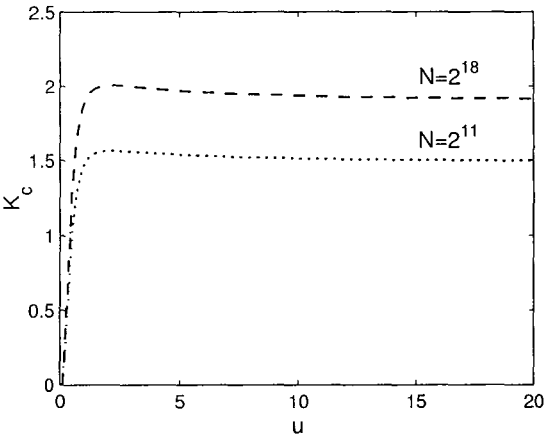


FIG. 3.13 – Variation de la constante K_c en fonction du paramètre u et de la taille du signal N.

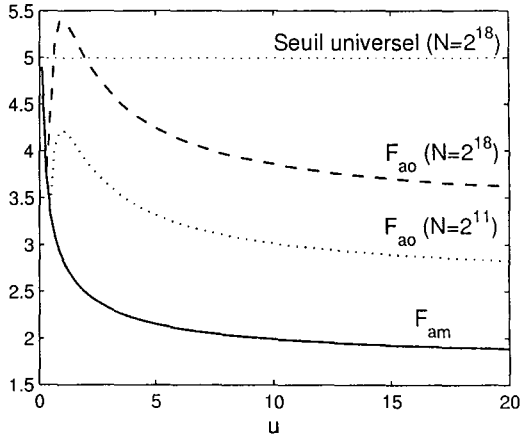


FIG. 3.14 – Variation de la constante multiplicative $F^j_{ao} = K^j_{ao} F^j_{am}$ en fonction du paramètre u des gaussiennes généralisées et de la taille du signal N.

avec

$$K^j_o = \max(1, K^j_c), \quad \text{et}$$
$$K^j_c = 2\sqrt{\frac{\alpha \log_e N}{3\beta\sqrt{2e}}} = \sqrt{\frac{2u \log_e N}{3\Gamma(\frac{1}{u})\sqrt{2e}}},$$

où N est la dimension du signal.

On vérifie facilement que $F^j_{ao} = \sqrt{2 \log_e N}$ pour les distributions gaussiennes ($u = 2$). L'expression de K^j_o se justifie par le fait que la valeur de K^j_c est inférieure à 1 pour certaines valeurs de u (voir figure 3.13), et donc le paramètre F^j_a calculé à partir de K^j_c devient inférieur à F^j_{am} et la convergence de l'algorithme n'est pas assurée. Comme pour l'algorithme *VisuShrink*, le seuillage "optimal" dépend de la taille N du signal (voir figure 3.14).

Comme nous l'avons déjà remarqué précédemment (section 3.2.2), même si le coefficient multiplicateur F_a est identique pour notre méthode et celle de Donoho et Johnstone, l'es-

timation de l'écart-type est différente (équations (3.9) et (3.10))[31], donc les seuils calculés sont différents. De plus, dans notre algorithme, ce seuil est encore diminué par la descente de point fixe ;

3. **seuillage par hystérésis** : ce dernier type de seuillage est une combinaison séquentielle des deux méthodes précédentes. Cette méthode utilise aussi l'information de segmentation échelle par échelle, décrite dans la section suivante (3.6). Le déroulement de l'algorithme suit trois étapes, qui s'appliquent indépendamment à toutes les échelles de la décomposition :
   (a) on calcule les deux seuils précédents ("optimal" et minimal). Le seuillage "optimal" détecte les grands coefficients d'ondelette, ce qui minimise le nombre de fausses alarmes, alors que le seuillage minimal minimise la perte d'information ;
   (b) à partir de chaque coefficient retenu par le seuillage "optimal" (qui correspond aux événements importants) on calcule les instants de début et de fin de l'événement détecté (donc le vecteur binaire de segmentation, voir les pas 2 et 3 de l'algorithme de segmentation, section 3.6) ;
   (c) pour chaque intervalle temporel obtenu après la segmentation précédente, on sélectionne *tous* les coefficients d'ondelette supérieurs au seuil minimal correspondants (dont les ondelettes associées contribuent à la reconstruction de l'événement inclus dans l'intervalle).

Autrement dit, nous utilisons le seuillage "optimal" pour détecter et segmenter rigoureusement les événements importants et minimiser les fausses alarmes, et le seuillage minimal pour garder un maximum d'information possible sur ces événements et assurer leur reconstruction précise.

3.6 Segmentation

La méthode décrite précédemment – la décomposition en ondelettes suivie du débruitage, favorise également une segmentation simple du signal[32]. En effet, les portions du signal éliminées par débruitage sont mises à 0, donc identifier un événement revient à détecter les parties non nulles.

Cependant, un des avantages de la décomposition en ondelettes est l'information conjointe temps-fréquence qu'elle offre pour décrire les événements (les sons abdominaux). L'information liée à la durée et à la position temporelle d'un événement est incorporée dans sa représentation en ondelettes : la reconstruction du signal débruité ne semble pas judicieuse à cette étape, et on préfère appliquer une segmentation directe sur le vecteur des coefficients d'ondelettes [87]. Plus précisément, on doit calculer la contribution de chaque ondelette dans le signal reconstruit. Une méthode simple, qui prend en compte la structure de l'algorithme et les dimensions des filtres de décomposition-reconstruction, est la suivante :

1. pour chaque échelle de détails $j = 1 \dots M$, on cherche les instants de fin des ondelettes d'indice p : $d = p \cdot 2^j$ (pour l'échelle basse fréquence de l'approximation $j = M + 1$, $d = p \cdot 2^M$) ;
2. on calcule les tailles des ondelettes respectives $T_j = (L-1) \cdot (2^j - 1) + 1$, ce qui permet de trouver leurs instants de début : $u = d - T_j + 1$, $j = 1 \dots M$, ($u = d - (L-1) \cdot (2^M - 1)$

[31] Dans notre interprétation, le point fixe de la fonction (3.17) peut coïncider avec le critère d'arrêt (3.14) seulement si l'écart-type est calculé à partir des coefficients (3.10). Ces deux valeurs sont proches quand les coefficients du signal informatif ne modifient pas l'écart-type, ou très peu, c'est-à-dire quand ils sont plutôt rares.

[32] La méthode présentée ici est valable pour une décomposition en ondelettes d'une taille égale à celle du signal, ce qui équivaut à une prise en compte des effets de bord par la périodisation du signal.

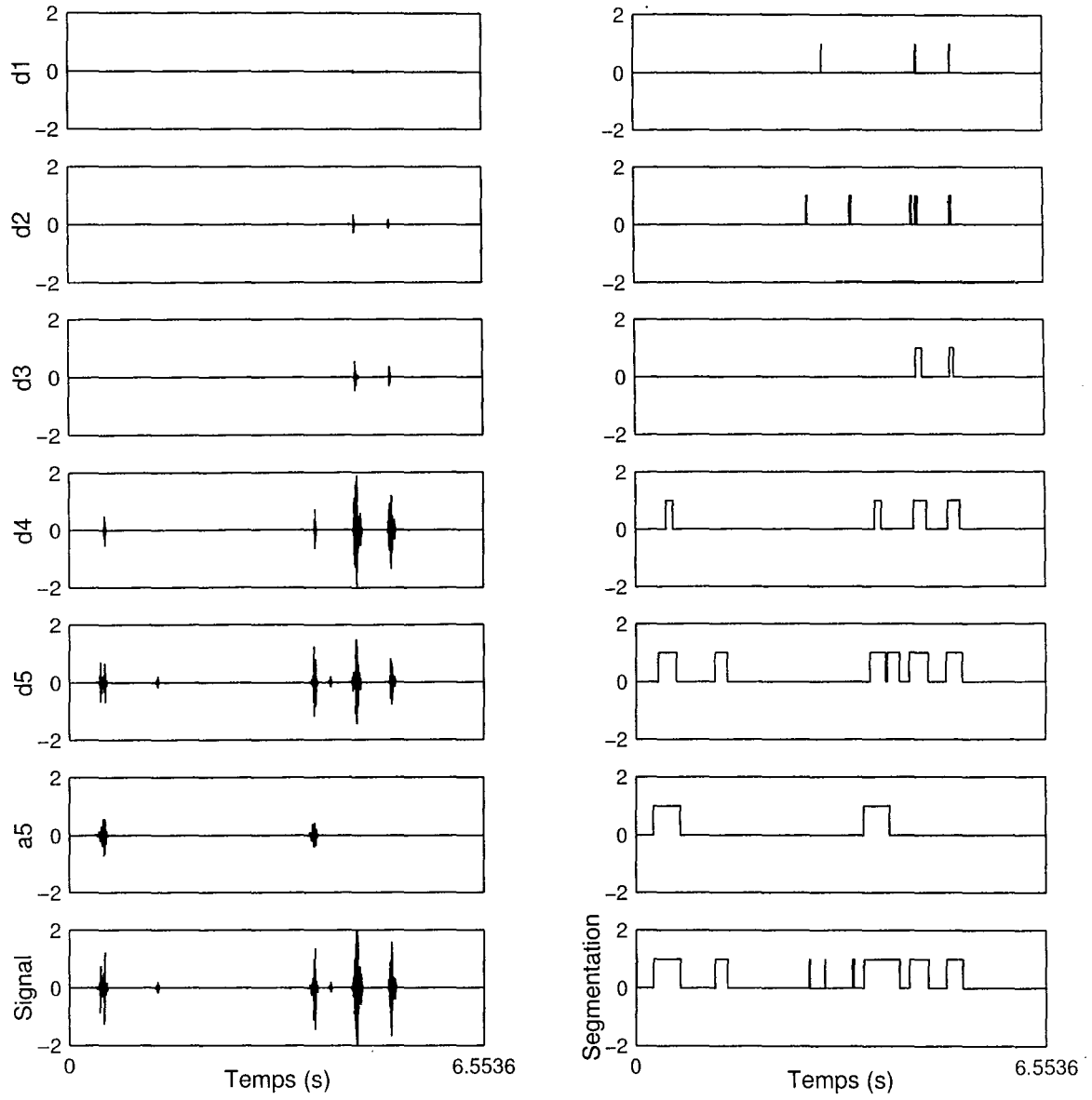


FIG. 3.15 – Illustration de la méthode de segmentation échelle par échelle : la segmentation des détails d'ondelettes est sommée pour créer le vecteur binaire de segmentation final (colonne de droite). La colonne de gauche représente les reconstructions des échelles non-nulles et leur somme, c'est-à-dire le signal débruité.

pour l'échelle $j = M + 1$), où L désigne la longueur de la réponse impulsionnelle des filtres (voir [71] pour les expressions des filtres associés aux ondelettes de Daubechies) ;

3. on crée, pour chaque échelle, un vecteur binaire (1 entre les instants de début et de fin, 0 ailleurs) ;

4. on applique un OU logique entre les vecteurs binaires correspondant aux échelles de la décomposition, pour créer un vecteur binaire de segmentation du signal débruité (voir figure 3.15).

3.7 Validation

Dans cette section, les résultats obtenus avec les trois versions de notre algorithme de débruitage décrit précédemment sont présentés et comparés aux résultats obtenus dans les mêmes conditions avec l'algorithme WTST-NST de Hadjileontiadis *et al.* et l'algorithme de débruitage *VisuShrink* de Donoho et Johnstone [37, 38]. Pour apprécier les performances de chacun des algorithmes, ils sont appliqués sur des signaux simulés.

3.7.1 Signaux considérés

Pour évaluer les performances des algorithmes, nous avons utilisé plusieurs combinaisons de signaux de test s et de bruits additifs r simulés.

Trois types de **signaux** test ont été employés :

1. signal nul ;
2. signaux artificiels type de Donoho et Johnstone : *Blocks*, *Bumps*, *HeaviSine* et *Doppler*, afin d'évaluer les différences de performance de débruitage[33] (figure 3.16) ;
3. signaux simulés proches de notre application. On rappelle que, dans le cadre de l'analyse des sons physiologiques, les signaux d'intérêt sont caractérisés par des événements isolés et courts, clairement séparés et avec un contenu fréquentiel relativement pauvre (la plupart des sons ont des fréquences entre 80 et 1000 Hz., même si certains auteurs signalent de rares événements de fréquences jusqu'à 2500 Hz. [15, 18, 41, 116]). Les signaux sont enregistrés avec des stéthoscopes, qui filtrent à la fois les sons abdominaux et le bruit. Pour pouvoir faire une analyse significative du point de vue médical, les enregistrements doivent être effectués sur de longues durées, et éventuellement être découpés en fenêtres d'observation plus courtes pour le traitement. Néanmoins, la longueur minimale de la fenêtre de traitement doit être adaptée à la profondeur choisie de la décomposition en ondelettes, pour assurer une estimation correcte de l'écart-type.

   Dans ce cadre, pour se placer dans des conditions qui couvrent notre domaine d'application, les tests ont été conduits sur une batterie de signaux simulés (échantillonnés à 5000 Hz. et ayant 2^{18} points, soit une durée d'environ 52 s.) composés d'une succession aléatoire d'événements (5, 10 ou 50 par minute). Pour conserver l'aspect impulsionnel des sons abdominaux, les événements générés sont des sinusoïdes de fréquences aléatoires comprises entre 80 et 1000 Hz. et d'amplitudes comprises entre 3 et 10, modulées par des exponentielles ou par des fenêtres de type Hanning de longueur aléatoire, entre 20 ms. et 1 seconde (figure 3.17).

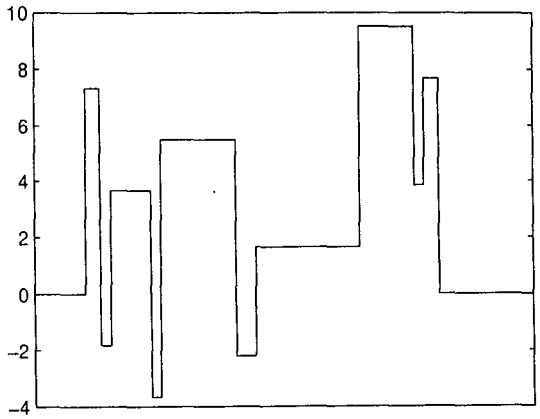
   La taille des signaux artificiels de Donoho et Johnstone (*Blocks*, *Bumps*, *HeaviSine* et *Doppler*) est de 2048 points, comme indiqué par les auteurs.

Nous avons généré plusieurs types de **bruits aléatoires**, superposés ensuite aux signaux d'intérêt. Nous avons fait varier plusieurs paramètres :

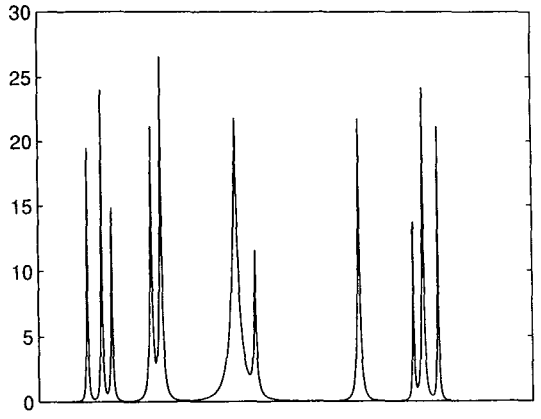
- la distribution du bruit (laplacienne, uniforme et gaussienne) ;
- la distribution fréquentielle (bruit blanc, bruit coloré proche du bruit réel – figure 3.18) ;
- la puissance (écarts-types de 1/2, 1 et 2).

La méthode de génération du bruit mérite quelques commentaires. En fait, dans une première étape, nous avons généré les bruits blancs comme des signaux aléatoires ayant les distributions et les écarts-types mentionnés. Le deuxième pas a été la génération des bruits colorés, par filtrage. Dans ce dernier cas, le filtrage change la nature du signal d'entrée : .

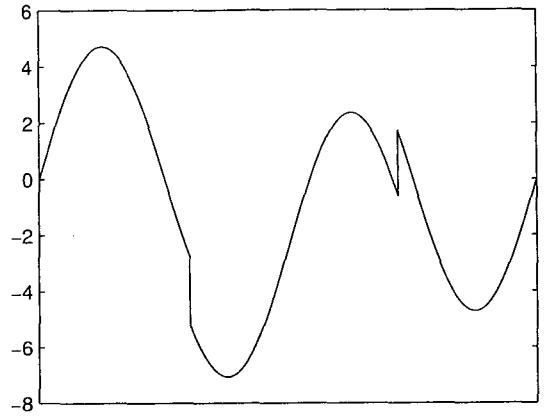
[33] Voir [38] pour les expressions des signaux.



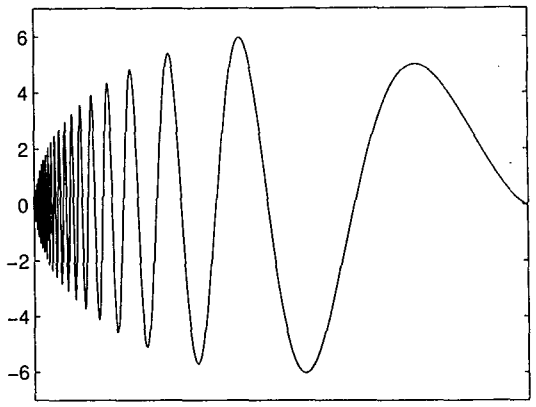
(a) Blocks



(b) Bumps



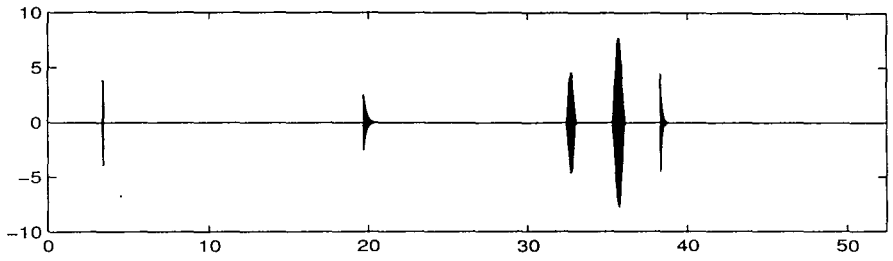
(c) HeaviSine



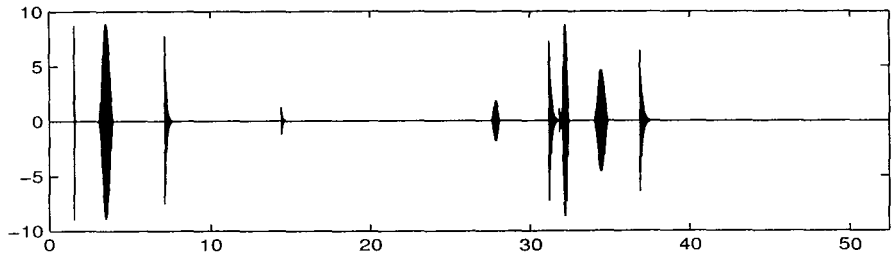
(d) Doppler

FIG. 3.16 – Signaux de test artificiels de Donoho et Johnstone [38].

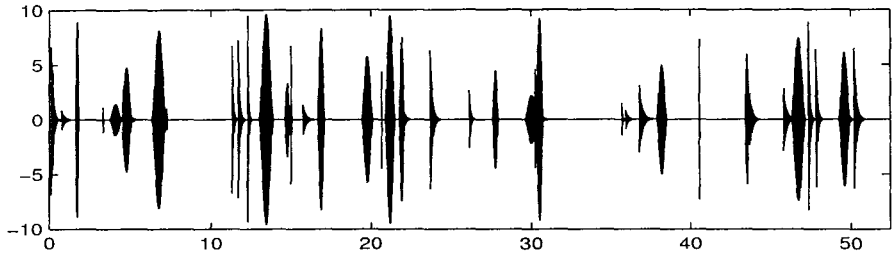
- la puissance du bruit coloré obtenu à la sortie du filtre est modifiée. Nous avons donc renormalisé les bruits, pour obtenir les écarts-types mentionnés. La conséquence directe de ce calcul est la modification différente en fonction de l'échelle des valeurs des coefficients d'ondelette. Ainsi, les basses fréquences du bruit coloré se retrouvent amplifiées par rapport aux basses fréquences du bruit blanc, ce qui aura pour conséquence la modification des seuils correspondants et donc de la qualité du débruitage (un seuil plus haut peut conduire à moins de fausses alarmes, voire à des non-détections) ;
- la distribution du bruit coloré est différente. Effectivement, la convolution avec un filtre revient à calculer chaque échantillon du signal de sortie (bruit coloré) comme une somme pondérée des échantillons du signal d'entrée (bruit blanc aléatoire). Conformément au théorème centrale-limite (généralisation de Lindeberg) [100], la distribution d'une somme de variables aléatoires indépendantes tend vers une distribution gaussienne. C'est notamment notre cas, comme nous pouvons le remarquer dans la figure 3.19.
  Cependant, les performances des algorithmes de débruitage varient en fonction de la dis-



(a) 5 événements



(b) 10 événements



(c) 50 événements

FIG. 3.17 – Signaux simulés proches des signaux réels de phonoentérographie. L'axe des abscisses représente le temps, en secondes.

tribution du bruit, surtout pour le bruit blanc, mais aussi, faiblement, pour le bruit coloré. Nous avons donc choisi de tester les algorithmes avec tous ces types de bruit, sachant que l'interprétation des résultats doit tenir compte des remarques précédentes.

3.7.2 Critères de performance

Nous avons utilisé plusieurs critères pour évaluer et comparer les performances des algorithmes de débruitage :

- l'erreur quadratique moyenne entre le signal original $\boldsymbol{s}$ et le signal débruité $\boldsymbol{c}$ [37, 71], qui est le critère le plus utilisé dans l'appréciation quantitative des performances des algorithmes :

$$EQM = E\left[(\boldsymbol{s} - \boldsymbol{c})^2\right] \ ; \tag{3.32}$$

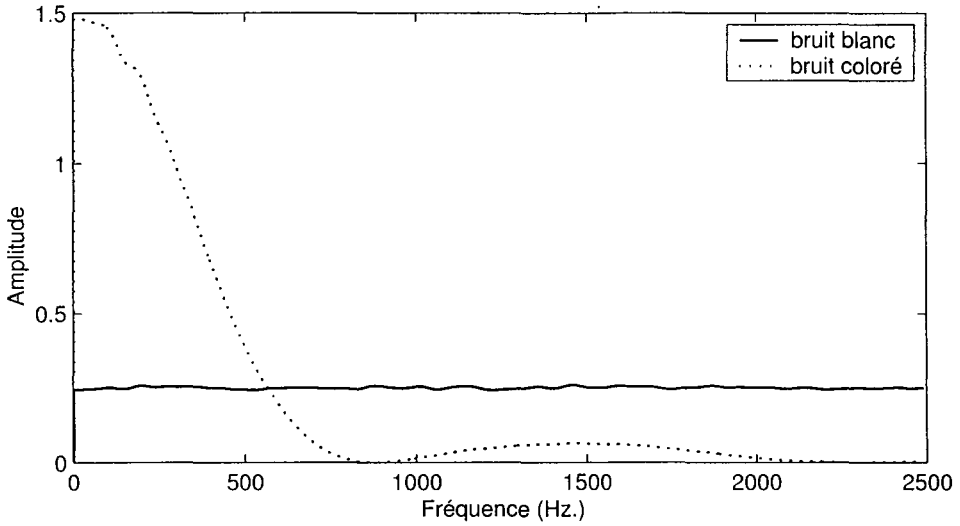


FIG. 3.18 – Spectres du bruit blanc et du bruit coloré simulés, de distribution laplacienne et écart-type $\sigma = 0,5$, utilisés pour tester les algorithmes de débruitage ($f_e = 5000$ Hz.).

- l'erreur de segmentation, définie à partir de l'évaluation de la durée totale du signal informatif (somme des durées des événements) rapportée à la durée du signal mesuré. Un signal creux pouvant être caractérisé par le rapport entre la durée totale de sa partie informative et sa longueur, nous calculons et comparons ce rapport pour le signal d'origine $\boldsymbol{s}$ et le signal débruité $\boldsymbol{c}$:

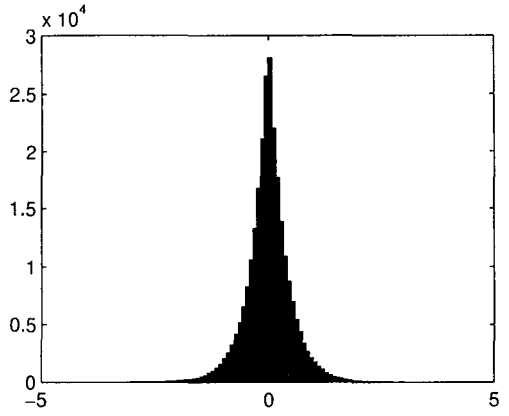
$$ES = \frac{\frac{D_{ev,c}}{D_c} - \frac{D_{ev,s}}{D_s}}{\frac{D_{ev,s}}{D_s}}, \tag{3.33}$$

  où $D_{ev,c}$, D_c, $D_{ev,s}$ et D_s sont la durée totale de tous les événements et la durée totale du signal débruité $\boldsymbol{c}$, respectivement du signal d'origine $\boldsymbol{s}$. La valeur positive ou négative de ES donne une indication sur la surévaluation ou la sous-évaluation de la durée totale des événements par rapport à la durée totale du signal, alors que la normalisation par rapport au signal d'origine nous permet d'obtenir une valeur relative au signal étudié ;
- l'erreur de détection, définie comme la différence normalisée entre le nombre d'événements détectés dans le signal débruité et le nombre d'événements dans le signal d'origine :

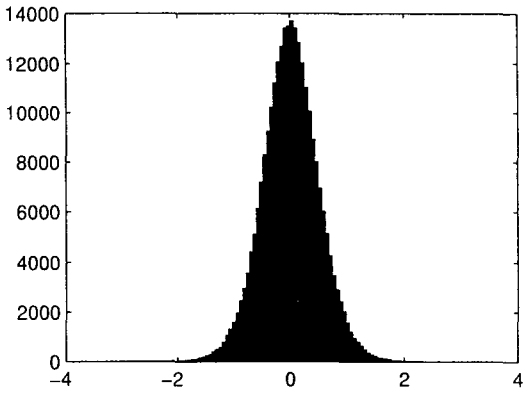
$$ED = \frac{N_{ev,c} - N_{ev,s}}{N_{ev,s}}, \tag{3.34}$$

  où $N_{ev,c}$ et $N_{ev,s}$ représentent respectivement le nombre d'événements dans le signal débruité et dans le signal d'origine. La valeur positive ou négative de ED donne une indication sur l'existence des fausses alarmes ou des non-détections.

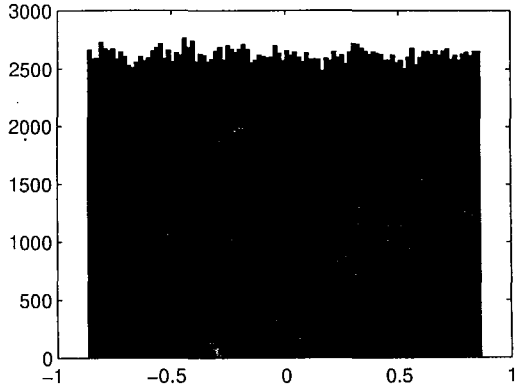
Les deux derniers critères, liés à la segmentation du signal, fournissent plutôt une évaluation qualitative, visuelle, du débruitage. Nous les introduisons comme des mesures supplémentaires de performance car nous considérons, comme Donoho et Johnstone [38], que l'aspect non-bruité d'un signal (*noise free*) est important pour son traitement et son interprétation ultérieurs. En effet, l'information fournie par ces deux critères est particulièrement importante car elle est la seule qu'on peut confronter à l'expertise médicale.



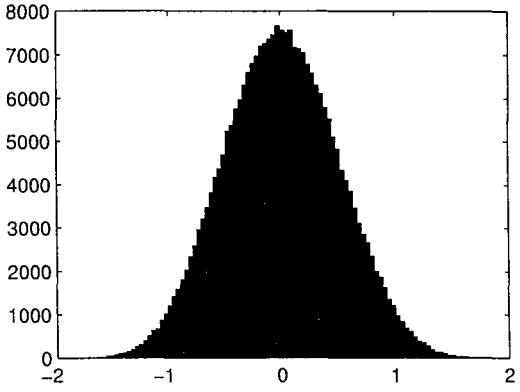
(a) bruit laplacien blanc ($\kappa = 5,93$)



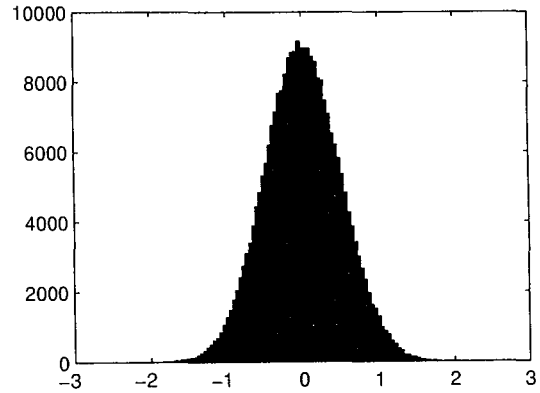
(b) bruit laplacien filtré ($\kappa = 3,56$)



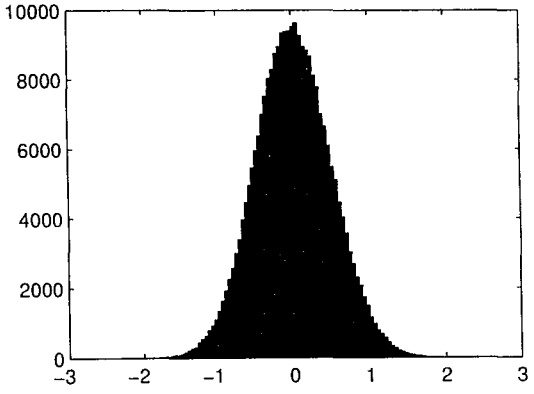
(c) bruit uniforme blanc ($\kappa = 1,80$)



(d) bruit uniforme filtré ($\kappa = 2,74$)



(e) bruit gaussien blanc ($\kappa = 3,00$)



(f) bruit gaussien filtré ($\kappa = 3,01$)

FIG. 3.19 – Histogrammes des bruits additifs utilisés pour la validation, avant filtrage et après filtrage. Entre parenthèses, nous donnons les valeurs empiriques du kurtosis (κ), qui permettent d'apprécier la gaussiannité d'une distribution ($\kappa = 3$ pour une distribution gaussienne).

3.7.3 Algorithmes testés

Les trois versions de notre algorithme (section 3.5.2) ont été comparées avec :
- l'algorithme itératif WTST-NST de Hadjileontiadis *et al.* ($F_a = 3$) [54, 55] ;
- l'algorithme *VisuShrink* de Donoho et Johnstone, avec l'estimation unique de l'écart-type sur la dernière échelle de la décomposition ;
- l'algorithme *VisuShrink* de Donoho et Johnstone, avec l'estimation de l'écart-type sur chaque échelle, pour considérer un bruit coloré (*VisuShrink-ME*) [37, 38, 71].

Une remarque importante doit être faite. Dans la version initiale de Donoho et Johnstone de l'algorithme *VisuShrink*, le seuillage des coefficients est appliqué seulement sur les échelles correspondant aux détails de la décomposition en ondelettes (les hautes fréquences), alors que les coefficients de l'approximation de la plus basse bande fréquentielle restent inchangés. Nous avons modifié la méthode, en imposant un traitement identique pour la première échelle, ce qui nous permet d'obtenir en même temps une segmentation du signal. En revanche, cette modification a des conséquences importantes sur le débruitage des signaux artificiels de Donoho et Johnstone [38], comme nous allons le montrer par la suite.

3.7.4 Base d'ondelettes

La base de décomposition-reconstruction en ondelettes choisie dans le même cadre applicatif par Hadjileontiadis *et al.* est l'ondelette de Daubechies no. 4, mais les tests réalisés sur plusieurs bases de la même famille ne montrent pas de modifications significatives des résultats [55]. Nous avons conservé la même approche, en testant plusieurs ondelettes (Daubechies et Symlets, d'ordres 4 à 23) et en comparant les résultats selon les trois critères précédents. L'algorithme de débruitage utilisé a peu d'importance pour la comparaison des performances en fonction du type d'ondelette, donc nous avons choisi le débruitage par hystérésis. Les signaux et les bruits additifs utilisés ont été générés comme décrit dans la section 3.7.1.

Les différences entre les résultats obtenus sont peu importantes, mais confirment en général l'utilité de l'adéquation de l'ondelette avec le signal traité. Globalement, les Daubechies se comportent mieux que les Symlets, probablement à cause du caractère asymétrique des signaux générés (voir annexe A pour une présentation des ondelettes testées). L'erreur quadratique moyenne (EQM) et l'erreur de détection (ED) diminuent avec l'ordre, ce qui s'explique par la plus grande décorrélation entre ces ondelettes et le bruit additif : moins de fausses alarmes sont extraites du bruit. En revanche, un ordre important est inapproprié pour les signaux de courte durée, car le support temporel est aussi plus grand et donc la segmentation est dégradée (erreur de segmentation ES plus importante).

Sur des signaux réels le comportement est analogue. Selon notre expérience, l'ondelette Daubechies no. 4 utilisée par Hadjileontiadis *et al.* détecte un nombre trop grand d'événements rapporté à l'expertise médicale (plus de 200 par minute). Ceci peut être expliqué par deux observations, liées à la taille réduite du support de l'ondelette :
- le morcellement de certains événements en plusieurs micro-événements ;
- l'extraction de fausses alarmes du signal mesuré[34].

Comme pour les signaux simulés, le nombre d'événements détectés diminue avec l'ordre de l'ondelette, à la fois par concaténation et par la réduction des fausses alarmes. En revanche, la durée totale des sons détectés augmente. L'ondelette Daubechies no. 9 nous a semblé être un bon compromis, mais ce choix est subjectif et discutable. Cependant, nous estimons que le changement de la base de décomposition apporte peu de différences sur les résultats du débruitage.

[34] En revanche, la durée totale des sons abdominaux est plus petite et donc le signal plus creux.

Les valeurs des critères de performance sont influencés aussi par la profondeur de la décomposition en ondelettes. Nous avons testé deux profondeurs M qui ont une signification physique :

- $M_1 = 7$ ($j = 1 \ldots 8$ échelles de décomposition) : le filtre passe-bas correspondant à la fonction d'échelle dans l'algorithme de Mallat coupe approximativement à 20 Hz. (fréquence minimale audible) et toute décomposition ultérieure est non-informative pour les signaux sonores ;
- $M_2 = 5$ ($j = 1 \ldots 6$ échelles de décomposition) : le filtre passe-bas correspondant à la fonction d'échelle coupe approximativement à 80 Hz. (fréquence minimale des sons abdominaux selon la bibliographie et selon notre expérience).

Les résultats des tests montrent que l'augmentation de la profondeur a un effet proche de l'augmentation de l'ordre : une profondeur accrue implique des échelles d'ondelettes de plus basse fréquence et de support plus grand, donc elle favorise la concaténation des micro-événements mais elle perturbe la segmentation. Notre choix s'est porté sur $M = 5$, mais la différence est peu importante, voire nulle pour certains signaux réels.

3.7.5 Résultats

Les tests se sont déroulés en trois étapes, correspondant aux types des signaux générés décrits précédemment (section 3.7.1).

Test 1. Le premier test cherche à étudier la capacité des algorithmes à reconnaître et éliminer le bruit. Nous avons donc considéré un signal de test de longueur $N = 2^{18}$ points composé uniquement de bruit, dont on fait varier la distribution et la puissance. Comme dans ce cas seul le critère de l'erreur quadratique moyenne (3.32) peut être appliqué (le nombre d'événements dans le signal d'origine et donc leur durée totale sont nuls), nous présentons dans le tableau 3.1 les valeurs de l'EQM pour des bruits blancs et colorés de différentes distributions. Comme nous l'avons montré précédemment (section 3.7.1), les distributions du bruit coloré obtenu par filtrage sont différentes de celles du bruit blanc. Cependant, dans le tableau 3.1 (comme dans tout le développement ultérieur) nous avons gardé les dénominations des distributions d'origine pour le bruit coloré, sous-entendant leur modification par filtrage.

Algorithme	Bruit blanc			Bruit coloré		
	Laplacien	Uniforme	Gaussien	Laplacien	Uniforme	Gaussien
"optimal"	0,0021	0	0	0,0002	0	0,0000
"minimal"	0,3897	0,1969	0,2714	0,2969	0,2531	0,2718
hystérésis	0,0022	0	0	0,0004	0	0,0000
WTST-NST	0,1677	0,0239	0,0604	0,0742	0,0519	0,0615
VisuShrink	0,0083	0	0	1,3439	1,2908	1,3168
VisuShrink-ME	0,0076	0	0	0,0005	0	0,0000

TAB. 3.1 – EQM pour $s = 0$ et différentes distributions de bruit blanc et coloré, toutes puissances confondues.

En analysant les résultats pour le bruit blanc, on peut remarquer les bonnes performances des algorithmes itératifs "optimal" et par hystérésis, ainsi que des algorithmes directs de type *VisuShrink* de Donoho et Johnstone. Les faibles différences de performance, dues à un seuillage plus fort pour nos algorithmes, sont essentiellement provoquées par la méthode d'estimation de l'écart-type (voir section 3.2.2). Comme attendu, l'algorithme "minimal" extrait trop d'information du bruit, comme l'algorithme WTST-NST de Hadjileontiadis *et al.*

Par rapport à la distribution de probabilité du bruit, les performances de tous les algorithmes sont inférieures pour la distribution laplacienne et meilleures pour l'uniforme. Dans le cas du bruit laplacien, on remarque les performances significativement supérieures des algorithmes itératifs "optimal" et par hystérésis face aux algorithmes directs de type *VisuShrink*. Ce constat était attendu : selon la figure 3.14, la constante multiplicative F_{ao} et donc les seuils des algorithmes "optimal" et par hystérésis sont supérieurs au seuil universel $\sqrt{2\log_e N}$ de *VisuShrink*, donc le débruitage est plus important.

Les mêmes constats sont valables pour le bruit coloré. Évidemment, l'algorithme *VisuShrink* sans estimation de l'écart-type échelle par échelle n'est pas adapté à l'élimination de ce type de bruit. Il est intéressant de noter que les performances de tous les algorithmes se rapprochent des valeurs obtenues dans le cas du bruit gaussien, ce qui confirme l'observation sur la tendance des distributions du bruit filtré vers des distributions gaussiennes. Néanmoins, les différences existent et le bruit, à l'origine, laplacien reste le plus difficile à éliminer, alors que le bruit uniforme semble le plus facilement rejetable.

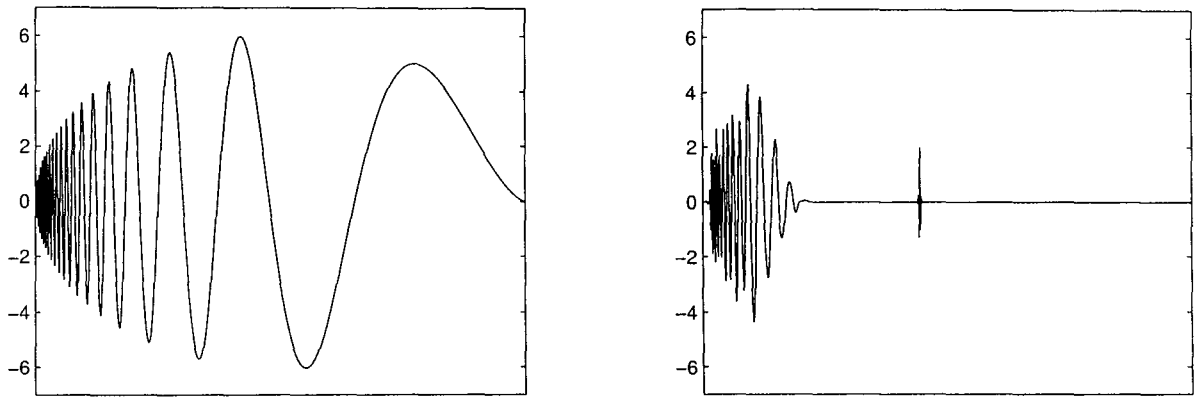
Enfin, nous avons testé les mêmes algorithmes en modifiant uniquement la longueur du signal ($N = 2^{21}$ points). Les résultats, qui ne seront pas détaillés ici, sont légèrement meilleurs, comme on pouvait s'y attendre : les algorithmes qui proposent un seuil dépendant de la taille du signal (tous, sauf les algorithmes "minimal" et WTST-NST) éliminent mieux le bruit (l'erreur quadratique moyenne tend vers 0 quand N tend vers l'infini), alors que les autres gardent des performances constantes.

Test 2. La deuxième étape de validation consiste en l'étude des performances des six algorithmes sur les signaux artificiels standard introduits par Donoho et Johnstone [38]. A ces signaux (de puissance égale), nous avons additionné les mêmes bruits que dans l'étape précédente de validation (Test 1). Nous avons utilisé comme critère de performance uniquement l'erreur quadratique moyenne EQM (3.32), car, à part le signal *Bumps*, les signaux de test ne sont pas des successions d'événements transitoires séparés par des zones de "silence", donc la segmentation et les critères dérivés (3.33) et (3.34) ne sont pas applicables. Nous présentons les performances des algorithmes testés dans le tableau 3.2.

Algorithme	Signal+bruit blanc				Signal+bruit coloré			
	Blocks	*Bumps*	*HSine*	*Doppler*	*Blocks*	*Bumps*	*HSine*	*Doppler*
"optimal"	20,14	12,95	13,23	12,03	20,18	13,25	13,23	12,27
"minimal"	20,07	3,10	13,44	12,15	20,02	7,89	13,37	12,17
hystérésis	20,14	12,93	13,23	12,00	20,17	13,24	13,23	12,23
WTST-NST	19,94	1,49	13,28	12,00	20,00	3,89	13,24	12,13
VisuShrink	0,39	0,42	0,15	0,21	1,45	1,46	1,42	1,43
VisuShrink-ME	19,93	0,84	13,24	12,02	20,07	3,04	13,23	12,19

TAB. 3.2 – EQM après débruitage pour les différents signaux de Donoho et Johnstone [38], toutes distributions et puissances de bruit confondues.

A la différence du premier test, pour les signaux artificiels de Donoho et Johnstone c'est l'algorithme *VisuShrink* avec estimation unique de l'écart-type qui se révèle le plus performant. L'explication réside dans les caractéristiques des signaux traités : ils ont des composantes informatives sur toute leur longueur. Autrement dit, le signal informatif n'est pas creux, il n'est pas rare dans le bruit. Cette caractéristique a des conséquences principalement dans les basses fréquences (voir figure 3.20) : comme les échelles correspondantes ont un nombre faible de coeffi-



(a) Signal Doppler d'origine

(b) Débruitage par l'algorithme par hystérésis

FIG. 3.20 – Exemple de débruitage inadéquat. On remarque l'élimination totale des basses fréquences.

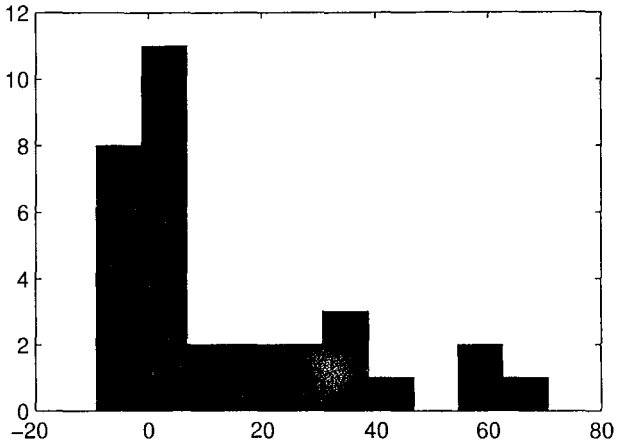


FIG. 3.21 – Histogramme des coefficients des ondelettes basses-fréquence du signal bruité *Bumps*.

cients d'ondelette, l'écart-type et donc le seuillage sont influencés d'une manière importante par les coefficients dus au signal informatif. Ces coefficients ne seront donc plus considérés comme des points aberrants (*outliers*) et ne seront plus détectés par seuillage. Ce constat explique d'ailleurs aussi les mauvaises performances de l'algorithme *VisuShrink-ME*[35].

Le signal *Bumps* mérite une remarque supplémentaire. En analysant les performances, on observe que deux des algorithmes itératifs ("minimal" et WTST-NST) ont des performances significativement meilleures que pour les autres signaux. Ceci s'explique par l'absence d'une composante fondamentale de basse-fréquence dans ce signal. Cependant, les performances des autres algorithmes itératifs, qui proposent un seuillage plus élevé, sont mauvaises. Outre les seuils plus importants, une cause possible est la distribution particulière des coefficients d'ondelette, toujours sur l'échelle correspondant à la plus basse-fréquence (figure 3.21). En effet, les algorithmes itératifs supposent une distribution centrée autour de 0, modélisable par une gaussienne généralisée, ce qui est loin d'être le cas pour le signal *Bumps*.

[35]Il s'agit des algorithmes modifiés par l'application du seuillage également dans les basses fréquences.

Test 3. La troisième étape de validation vise à évaluer les performances des algorithmes de débruitage sur des signaux simulés proches des signaux réels de notre application. Aux signaux de test générés avec la technique décrite dans la section 3.7.1, nous avons superposé les mêmes types de bruit que dans les étapes précédentes de validation. Nous présentons les performances des algorithmes selon les trois critères définis dans (3.32), (3.33) et (3.34). Comme pour les signaux artificiels de Donoho et Johnstone (tableau 3.2), nous présentons dans les tableaux 3.3, 3.4 et 3.5 les résultats par type de signal (5 év./minute, 10 év./minute, 50 év./minute) et de bruit (blanc ou coloré), sans détailler en fonction de la distribution ou de la puissance du bruit.

Algorithme	Signal+bruit blanc			Signal+bruit coloré			Moyenne
	5 év.	10 év.	50 év.	5 év.	10 év.	50 év.	générale
"optimal"	0,03	0,10	0,53	0,08	0,14	0,41	0,22
"minimal"	0,28	0,28	0,38	0,26	0,26	0,33	0,30
hystérésis	0,02	0,08	0,41	0,06	0,10	0,31	0,16
WTST-NST	0,10	0,13	0,28	0,10	0,12	0,25	0,16
VisuShrink	0,03	0,09	0,38	1,32	1,33	1,37	0,75
VisuShrink-ME	0,03	0,10	0,43	0,08	0,12	0,39	0,19

TAB. 3.3 – EQM après débruitage pour les différents signaux simulés (variation du nombre d'événements/minute), toutes distributions et puissances de bruit confondues.

Le critère classique d'évaluation pour les algorithmes de débruitage est l'erreur quadratique moyenne normalisée (3.32). Appliqué aux algorithmes et aux signaux décrits précédemment, il fournit les résultats présentés dans le tableau 3.3 : la troisième version de notre algorithme, le seuillage par hystérésis, se révèle la plus performante pour les signaux à faible nombre d'événements par minute (5 et 10, avec des pourcentages de présence d'événements de 5,8% et respectivement 10,4%). L'algorithme "optimal" a des performances légèrement inférieures aux algorithmes directs de type *VisuShrink*, sauf bien-sûr pour le bruit coloré, car l'algorithme *VisuShrink* classique n'est pas adapté. Enfin, l'algorithme WTST-NST et surtout l'algorithme "minimal" ont tendance à extraire trop d'information et à garder trop de bruit dans le signal débruité. Les bonnes performances de l'algorithme par hystérésis sont dues à la combinaison des deux seuillages proposés par les versions "optimale" et "minimale". En fait, grâce à l'algorithme "minimal", l'algorithme par hystérésis garde un maximum d'information autour des instants et fréquences significatives détectés par l'algorithme "optimal".

La situation change pour les signaux à grand nombre d'événements par minute (50, avec un pourcentage de présence d'événements de 38,9%). Les algorithmes proposant un seuillage fort perdent leur avantage :

- par leur construction, les algorithmes itératifs ("optimal" et par hystérésis) détectent les exceptions, et le caractère exceptionnel des événements leur est ôté par leur nombre important et donc par leur influence sur l'écart-type ;
- les algorithmes directs (*VisuShrink* et *VisuShrink-ME*), qui sont basés aussi sur l'écart-type, proposent un seuil trop important à cause d'une valeur trop élevée de celui-ci.

Dans ces conditions, si l'algorithme "minimal" reste toujours trop peu sélectif, l'algorithme WTST-NST de Hadjileontiadis *et al.* offre une solution *ad hoc*, surtout dans sa version rapide, améliorée par l'interprétation de point-fixe que nous avons proposée (section 3.4).

Concernant l'erreur d'évaluation de la durée totale du signal informatif ES (tableau 3.4), les algorithmes se comportent de deux manières différentes :

Algorithme	Signal+bruit blanc			Signal+bruit coloré			Moyenne
	5 év.	10 év.	50 év.	5 év.	10 év.	50 év.	générale
"optimal"	-0,40	-0,48	-0,54	-0,35	-0,27	-0,34	-0,40
"minimal"	13,42	6,68	0,73	12,71	6,32	0,64	6,75
hystérésis	-0,39	-0,48	-0,54	-0,34	-0,17	-0,34	-0,38
WTST-NST	5,50	2,69	0,25	5,01	2,56	0,30	2,72
VisuShrink	-0,37	-0,48	-0,53	x	x	x	x
VisuShrink-ME	-0,44	-0,53	-0,56	-0,40	-0,36	-0,41	-0,45

TAB. 3.4 – *ES* après débruitage pour les différents signaux simulés (variation du nombre d'événements/minute), toutes distributions et puissances de bruit confondues.

- les algorithmes itératifs à seuillage faible et indépendant de la taille du signal ("minimal" et WTST-NST) surévaluent cette durée (jusqu'à des valeurs 13 fois plus importantes, dans le cas de l'algorithme "minimal" et du signal à 5 événements par minute), ce qui suggère la présence importante de bruit dans le signal débruité ;
- les algorithmes à seuillage fort (dépendant de la taille du signal) sous-estiment la durée du signal informatif (jusqu'à moins de la moitié, pour les signaux à 50 et même à 10 événements par minute), ce qui indique que les parties de faible amplitude du signal informatif ont été considérées comme du bruit et éliminées.

Le classement des algorithmes est légèrement différent de celui en fonction de l'EQM, mais il respecte la distinction entre les signaux à faible nombre d'événements par minute (5 ou 10) et ceux à grand nombre d'événements (50). Ainsi, l'élimination du bruit blanc superposé aux signaux à faible nombre d'événements est meilleure avec l'algorithme *VisuShrink*. Cet algorithme est suivi par le débruitage par hystérésis et l'algorithme "optimal". En revanche, pour le bruit coloré, le seuillage par hystérésis fournit les meilleurs résultats, à l'exception du cas des signaux riches en événements.

Il est intéressant de noter que l'évaluation de la durée totale du signal informatif (liée à la segmentation) est meilleure dans le cas d'un bruit coloré. Ceci s'explique par la forte atténuation du bruit dans les hautes fréquences, ce qui conduit à un faible nombre de fausses alarmes sur les échelles correspondantes de la décomposition en ondelettes. Cette hypothèse est soutenue par les valeurs du troisième critère de performance utilisé (*ED*), lié au nombre d'événements détectés (tableau 3.5).

Algorithme	Signal+bruit blanc			Signal+bruit coloré			Moyenne
	5 év.	10 év.	50 év.	5 év.	10 év.	50 év.	générale
"optimal"	1,09	1,23	1,26	0,72	1,34	1,64	1,21
"minimal"	118,02	70,06	30,28	106,09	40,21	4,08	61,46
hystérésis	1,04	1,22	1,18	0,70	1,32	1,61	1,18
WTST-NST	101,59	59,07	17,62	87,96	49,87	14,52	55,10
VisuShrink	3,52	2,44	1,75	x	x	x	x
VisuShrink-ME	3,30	2,18	1,56	0,72	0,70	0,90	1,56

TAB. 3.5 – *ED* après débruitage pour les différents signaux simulés (variation du nombre d'événements/minute), toutes distributions et puissances de bruit confondues.

L'erreur de segmentation présentée dans le tableau 3.5 montre que le débruitage par hystérésis semble le plus performant dans le cas du bruit blanc, mais que l'algorithme *VisuShrink-ME* est meilleur pour la détection dans un bruit coloré, surtout si le nombre d'événements est important. L'algorithme WTST-NST, qui offrait de bonnes performances selon les autres critères pour ce dernier type de signaux, génère beaucoup trop de fausses alarmes.

Une image plus détaillée des performances comparatives des algorithmes de débruitage (en fonction de la distribution du bruit et de sa puissance) est donnée sur les figures B.1 à B.12, présentées dans l'annexe B. Les résultats généraux discutés précédemment se confirment, avec quelque nuances : si la supériorité de l'algorithme de débruitage par hystérésis est évidente pour un bruit additif laplacien, tant du point de vue de l'erreur quadratique (EQM) moyenne que des critères de segmentation (ES et ED), le choix du meilleur algorithme est plus problématique dans le cas du bruit gaussien. En effet, même si l'EQM de l'algorithme "par hystérésis" est plus petite que celle du *VisuShrink-ME*, la détection des événements est plus précise par ce dernier. Ces deux situations sont illustrées par les figures 3.22 et 3.23.

Implémentation informatique. Nous avons traité des signaux de 2^{18} points, à une fréquence d'échantillonnage de 5000 Hz. (ce qui correspond à un enregistrement d'une durée d'environ 52 secondes), sur un ordinateur de type PC Pentium IV 1,6 GHz., les algorithmes étant codés principalement en langage Matlab. Les temps d'exécution des différents algorithmes, en incluant la décomposition initiale et la reconstruction finale des signaux, sont présentés dans le tableau 3.6. Nous mentionnons également les temps de calcul de l'algorithme itératif original WTST-NST(o) de Hadjilontiadis *et al.* (avec décomposition et reconstruction à chaque itération).

Algorithme	Durée [s]
"optimal"	2,35
"minimal"	4,66
hystérésis	3,72
WTST-NST(pf)	4,13
VisuShrink	1,47
VisuShrink-ME	1,85
WTST-NST (o)	18,44

TAB. 3.6 – Comparaison des temps d'exécution des algorithmes. L'algorithme WTST-NST(o) est l'algorithme original de Hadjileontiadis *et al.* sans l'optimisation de point fixe, avec une valeur recommandée par les auteurs de la constante $\varepsilon = 0,00001$.

Nous observons donc un gain en temps de calcul évident grâce à l'interprétation comme algorithme de point fixe : le facteur de gain dépend du nombre d'itérations, qui lui même dépend de la valeur du seuil (liée aux caractéristiques du signal) et du paramètre ε qui conditionne le critère d'arrêt STC (éq. (3.12)). En appliquant cette nouvelle interprétation, l'approche itérative présente un temps d'exécution du même ordre de grandeur que l'approche directe de Donoho et Johnstone [37, 38]. Cependant, on observe des variations de durée entre les différents algorithmes :

- entre les méthodes "optimale" et "minimale", la différence est essentiellement due à la valeur du seuil et au nombre d'itérations vers le point fixe : une constante F_a plus élevée implique une fonction plus "abrupte" et un seuil plus grand ;
- même si on applique deux seuillages, l'algorithme par hystérésis est plus rapide que l'algorithme "minimal", car le premier seuillage est haut ("optimal"), l'information recherchée par le seuillage "minimal" étant limitée aux événements détectés par la première étape ;

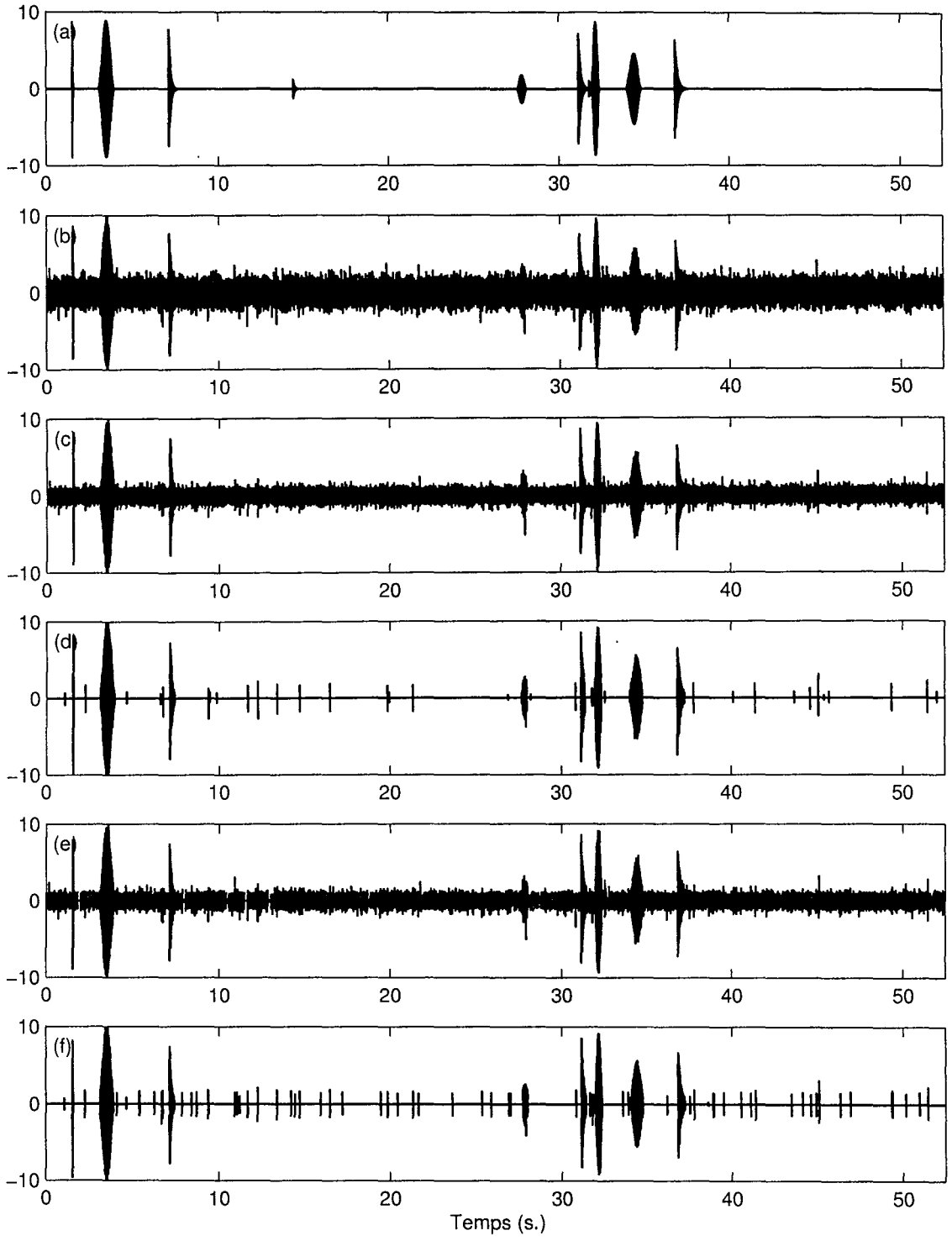


FIG. 3.22 – Résultats comparatifs du débruitage pour un bruit additif blanc, de distribution laplacienne : (a) signal d'origine ; (b) signal bruité ; signaux débruités par (c) algorithme "minimal", (d) algorithme par hystérésis, (e) WTST-NST, (f) *VisuShrink*. Le résultat fourni par l'algorithme "optimal" n'est pas présenté ici, car le signal est visuellement indistinct de (d). Le même argument est valable pour l'algorithme *VisuShrink-ME*, étant donné qu'il s'agit d'un bruit additif blanc.

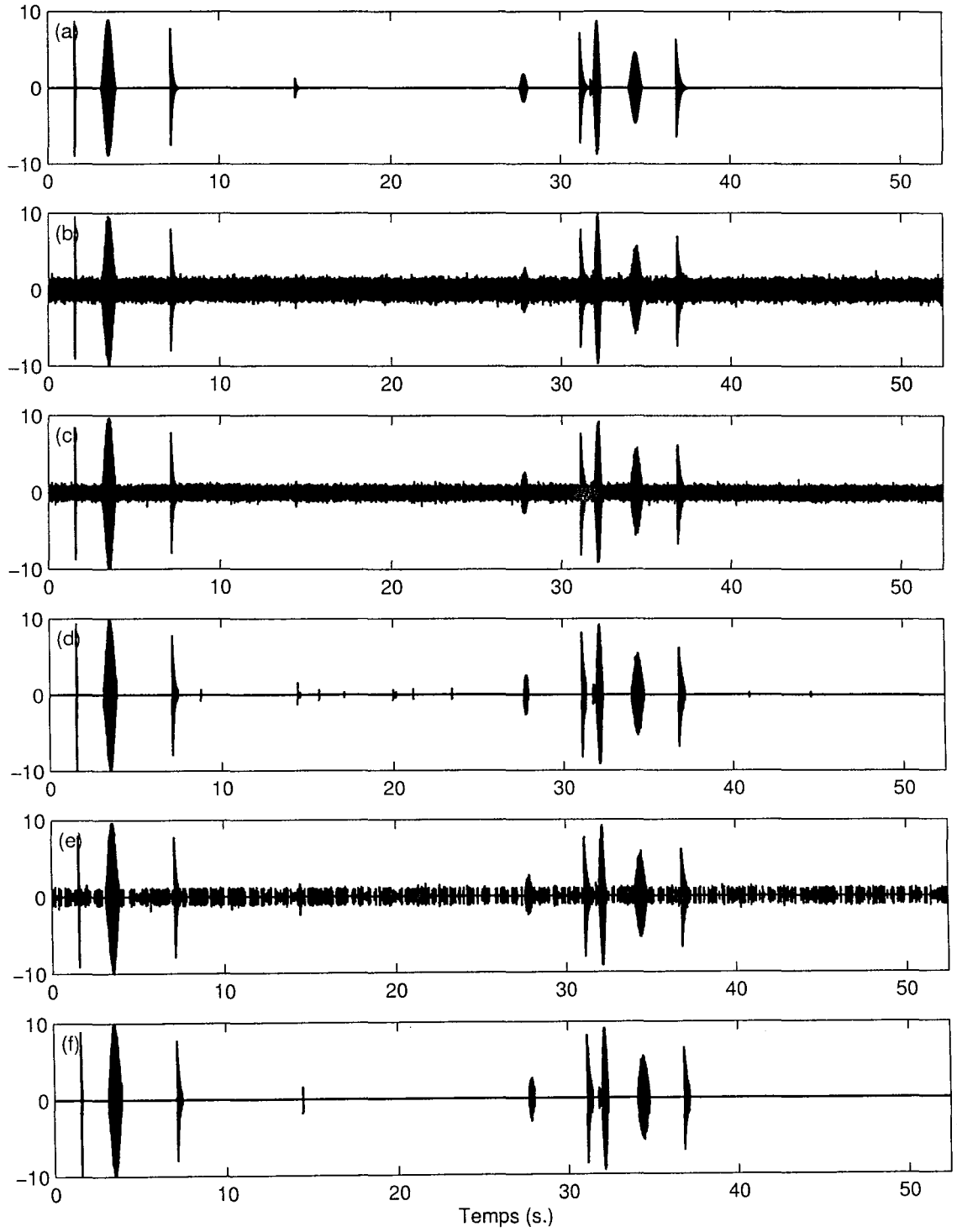


FIG. 3.23 – Résultats comparatifs du débruitage pour un bruit additif blanc, de distribution gaussienne : (a) signal d'origine ; (b) signal bruité ; signaux débruités par (c) algorithme "minimal", (d) algorithme par hystérésis, (e) WTST-NST, (f) *VisuShrink*. Le résultat fourni par l'algorithme "optimal" n'est pas présenté ici, car le signal est visuellement indistinct de (d). Le même argument est valable pour l'algorithme *VisuShrink-ME*, étant donné qu'il s'agit d'un bruit additif blanc.

- la différence entre nos algorithmes ("optimal", "minimal" et par hystérésis) et l'algorithme WTST-NST(pf) n'est pas significative, car la modélisation probabiliste de la distribution des coefficients par des gaussiennes généralisées, n'augmente pas le temps de calcul, grâce à l'implémentation en langage C[36].

3.8 Conclusion

De ce chapitre, nous pouvons conclure que l'extraction d'événements dans les signaux non-stationnaires bruités doit être précédée d'une réflexion sur les méthodes à mettre en œuvre, réflexion qui doit prendre en compte les éventuelles connaissances *a priori* sur le signal informatif et le bruit. L'approche basée sur la décomposition en ondelettes orthogonales nous semble appropriée :

- les ondelettes sont adaptées pour la détection des régimes transitoires dans les signaux ;
- l'orthogonalité assure un passage entre les représentations du signal qui conserve l'énergie ;
- le débruitage par ondelettes (qui peut être vu comme un filtrage non-linéaire [71]) assure en même temps une segmentation du signal ;
- enfin, par la représentation conjointe temps-fréquence, la décomposition en ondelettes prépare l'analyse des événements segmentés pour les étapes ultérieures.

L'analyse globale des performances des algorithmes de débruitage permet de choisir l'algorithme le mieux adapté à notre application. Nous considérons que, vu le faible niveau de connaissances sur la nature du bruit et du signal informatif des phonoentérogrammes, l'algorithme de débruitage par hystérésis est le plus approprié.

De plus, les deux seuils proposés par cet algorithme (calculés par la méthode itérative de point fixe présentée dans ce chapitre) semblent s'adapter mieux au contenu fréquentiel du signal informatif que les algorithmes directs de type *VisuShrink*. Par exemple, pour le signal simulé à 5 événements par minute, une quantité importante d'information est concentrée entre 80 et 160 Hz., bande fréquentielle correspondant à la cinquième échelle de la décomposition en ondelettes (voir figure 3.24).

Les seuils calculés échelle par échelle par plusieurs algorithmes de débruitage sont représentés sur les figures 3.25(a) et 3.25(b). On observe un seuil plus haut de l'algorithme "minimal" (par rapport à l'algorithme WTST-NST) dans la bande de fréquence mentionnée (ce qui s'explique par la densité plus grande du signal informatif dans cette bande), mais un seuil plus bas de l'algorithme "optimal" (rapporté à l'algorithme *VisuShrink-ME*). Nous considérons que cette capacité accrue de détection dans la bande fréquentielle d'intérêt est due à la nature itérative de l'algorithme, qui permet une réduction du niveau de seuillage adaptée à l'information contenue dans le signal.

Cependant, le succès de l'algorithme dépend des particularités du signal informatif et du bruit superposé. La caractéristique la plus importante semble être le nombre d'événements par unité de temps. On constate que les algorithmes de débruitage par hystérésis et "optimal" ne sont pas adaptés pour les signaux riches en événements (50 év./min.). En revanche, l'algorithme par hystérésis se comporte mieux que les autres, quel que soit le critère retenu, pour les signaux pauvres en événements (jusqu'à 10 év./min.). De plus, la supériorité de notre algorithme est incertaine dans le cas d'un bruit additif gaussien, alors qu'elle est beaucoup plus claire pour un bruit laplacien.

[36] Le code des algorithmes de modélisation de la distribution des coefficients par des lois gaussiennes généralisées est écrit par Daniel Schöbben et il est disponible à l'adresse Internet `http://www.esp.ele.tue.nl/onderzoek/daniels/gengauss/gengauss.html` [101].

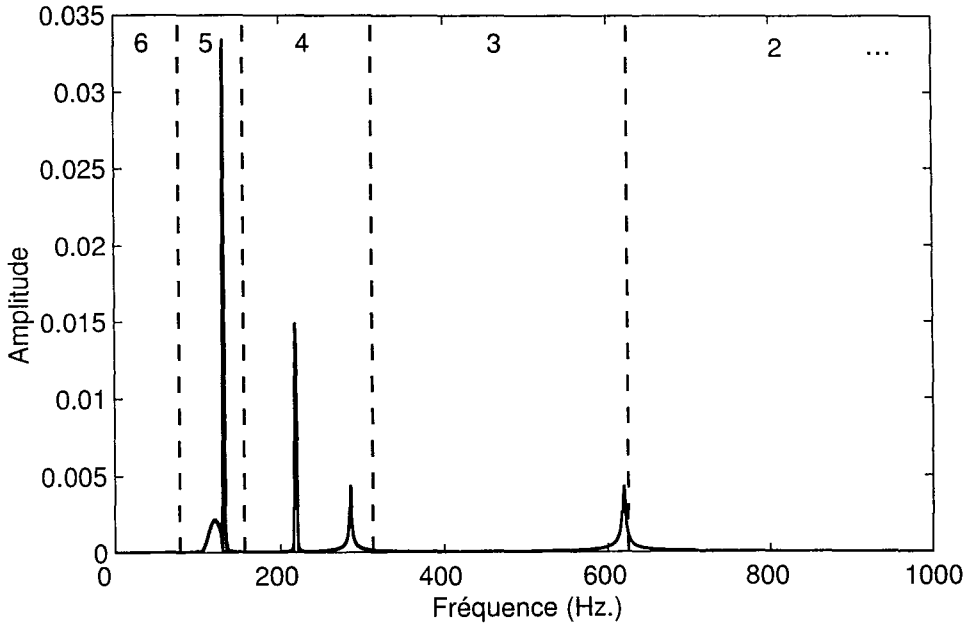
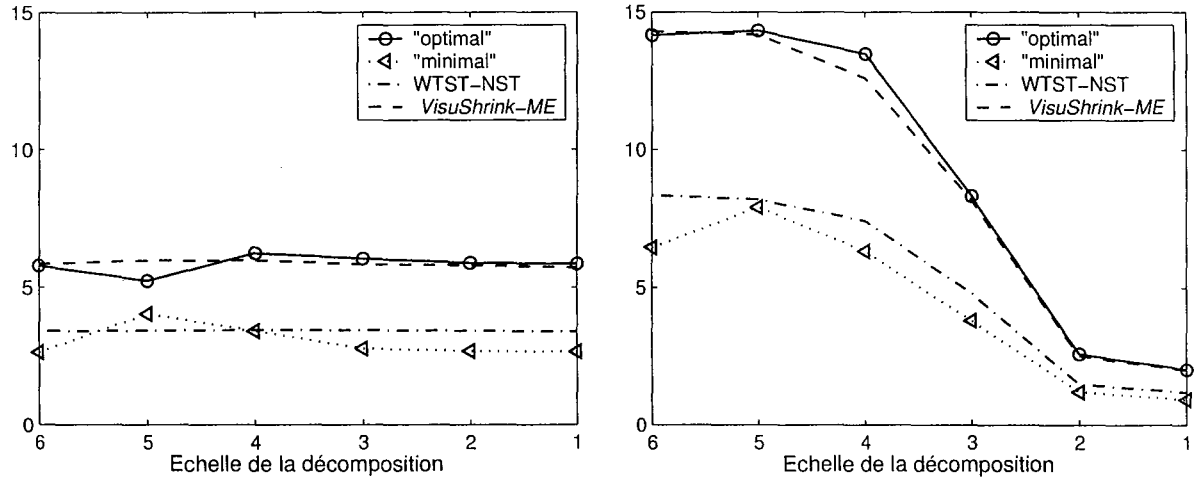


FIG. 3.24 – Exemple de spectre de signal simulé à 5 événements par minute, avec les échelles correspondantes aux bandes de fréquence délimitées par les lignes en pointillé.



(a) bruit blanc

(b) bruit coloré

FIG. 3.25 – Exemples de variations des seuils des différents algorithmes de débruitage par échelle. Les valeurs représentées ici sont des moyennes pour le signal à 5 événements par minute, toutes distributions et puissances de bruit confondues.

En outre, comme l'exemple du débruitage du signal artificiel *Bumps* nous l'a montré, l'algorithme est applicable seulement si les coefficients d'ondelette ont une distribution proche d'une gaussienne généralisée.

En effet, notre algorithme n'est pas strictement un algorithme de débruitage, mais plutôt un algorithme d'extraction d'événements, ou de débruitage et segmentation. Il est plus efficace quand le signal informatif est rare dans le bruit.

Dans notre application, notre préoccupation est de traiter les sons physiologiques sur lesquels nous n'avons que très peu de connaissances *a priori*, hormis leur non-stationnarité. Comme on peut observer sur les exemples de débruitage des signaux réels présentés figures 3.26 et 3.27, il est difficile de choisir entre l'algorithme direct *VisuShrink-ME* de Donoho et Johnstone (avec les modifications concernant la première échelle – voir section 3.7.1), et l'algorithme de débruitage itératif par hystérésis que nous avons développé à partir des travaux de Hadjileontiadis *et al.* dans notre domaine d'intérêt. Nous considérons que notre méthode s'adapte mieux aux variations des caractéristiques du signal et à son contenu fréquentiel. Dans un souci de généralité et étant donné que nos connaissances actuelles sur les sons abdominaux se limitent à quelques cas particuliers, nous avons choisi de l'utiliser en priorité. Cependant, un choix définitif devra être fait en fonction des résultats des étapes ultérieures d'analyse et interprétation des phonoentérogrammes, après des essais cliniques. De plus, nous ne pouvons pas exclure la nécessité d'un traitement ultérieur, destiné à rejeter une partie de l'information extraite par une des méthodes de débruitage et segmentation présentées dans ce chapitre, car trop de données peuvent nuire à une classification et à une interprétation pertinentes et applicables en routine clinique.

Enfin, nous tenons à noter que le domaine de la segmentation et du débruitage par ondelettes mérite d'être exploré davantage. En particulier, nous considérons que d'autres améliorations des algorithmes itératifs constituent une perspective de recherche intéressante.

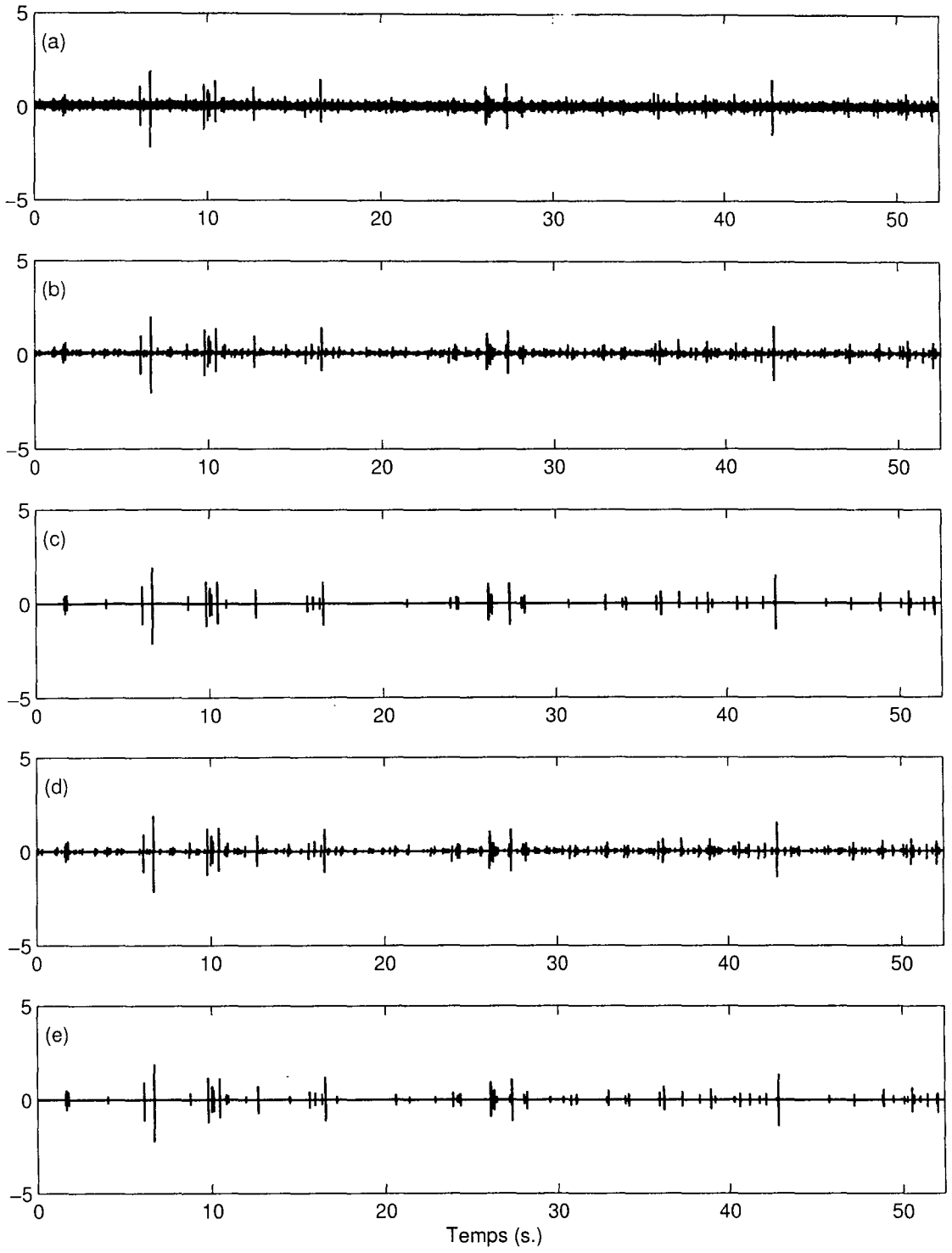


FIG. 3.26 – Résultats comparatifs du débruitage pour un premier signal réel d'environ 52 s. (entre parenthèses, nous donnons l'évaluation du pourcentage de la durée totale occupé par le signal informatif D et le nombre d'événements détectés N_{ev}) : (a) signal mesuré ; (b) algorithme "minimal" ($D = 100\%, N_{ev} = -$) ; (c) algorithme par hystérésis ($D = 7,9\%, N_{ev} = 76$) ; (d) WTST-NST ($D = 70,5\%, N_{ev} = 334$) ; (e) *VisuShrink-ME* ($D = 16,5\%, N_{ev} = 86$).

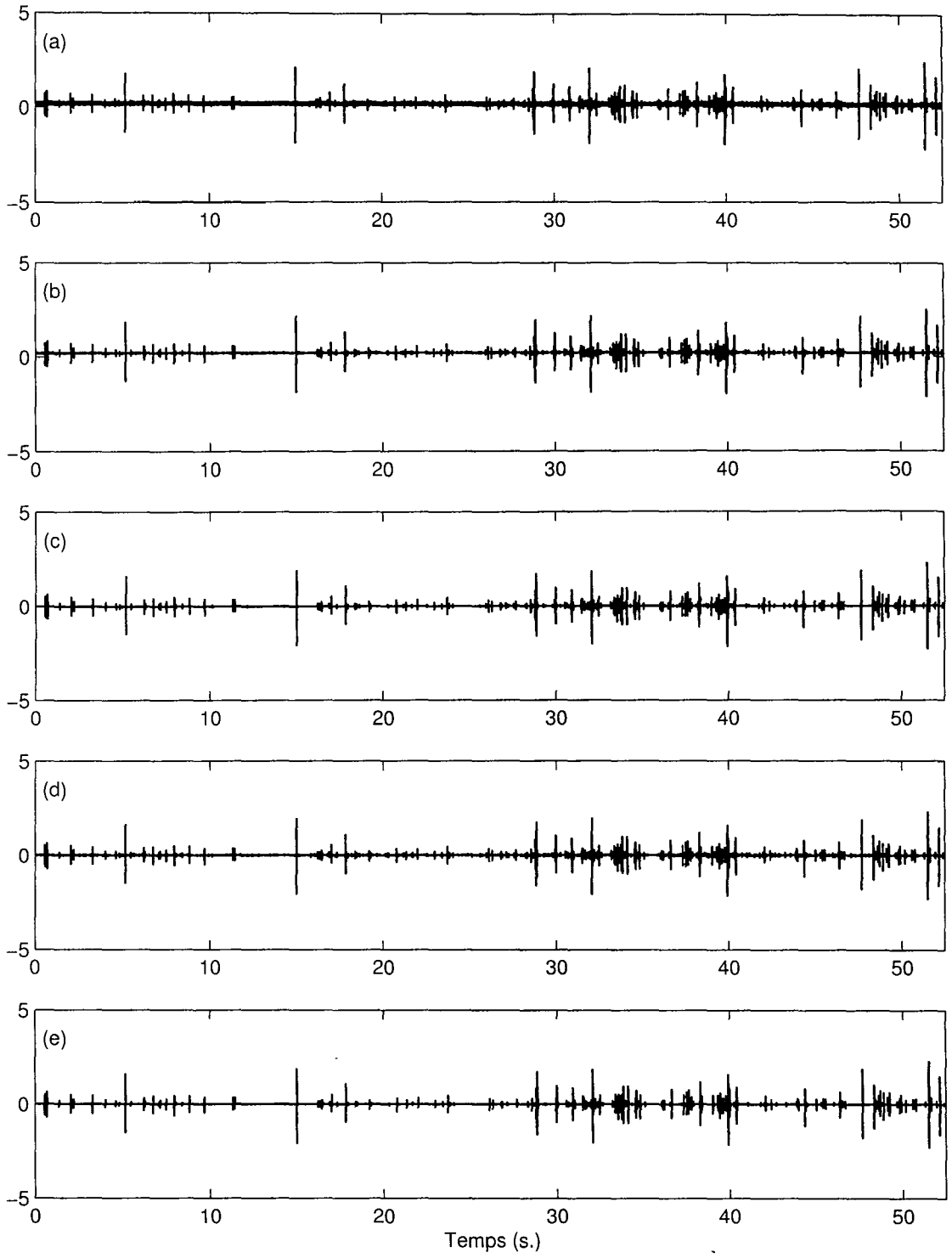


FIG. 3.27 – Résultats comparatifs du débruitage pour un deuxième signal réel d'environ 52 s. : (a) signal mesuré ; (b) algorithme "minimal" ($D = 100\%, N_{ev} = -$) ; (c) algorithme par hystérésis ($D = 22,2\%, N_{ev} = 191$) ; (d) WTST-NST ($D = 42,7\%, N_{ev} = 604$) ; (e) *VisuShrink-ME* ($D = 17,3\%, N_{ev} = 159$).



Segmentation et localisation

Le pré-traitement détaillé dans le chapitre précédent est une première étape essentielle vers l'objectif que nous nous sommes fixé. L'algorithme de détection et débruitage est relativement général et applicable à de nombreux types de signaux non-stationnaires[37]. Cependant, il ne prend pas en compte ni les connaissances *a priori* sur les caractéristiques fréquentielles et temporelles des signaux, ni la nature multi-voies des enregistrements. L'objectif de ce chapitre est l'intégration de ces informations dans le traitement des phonoentérogrammes, par le biais de la re-segmentation des signaux et de la localisation des sons abdominaux.

Nous favorisons donc une approche empirique, pragmatique, fondée sur les caractéristiques physiques des sons abdominaux présentées dans la section 2.1. La justification principale de cette démarche est la nécessité de l'intégration de l'expertise médicale dans le traitement des signaux, étant donné qu'elle est la seule source d'information et de validation pour le volet applicatif de nos travaux. Le premier objectif de ce chapitre est l'élimination des artéfacts, définis comme les événements qui ne correspondent pas aux connaissances *a priori* sur les sons abdominaux. Il s'agit, par exemple, des fausses alarmes restantes après le débruitage, des bruits "aigus" dus aux micro-frottements entre la tête de stéthoscope et la peau ou des bruits "graves" dus aux chocs accidentels ou aux mouvements du patient.

Le protocole de mesure défini précédemment (section 2.2.1) permet l'enregistrement simultané de plusieurs signaux. Les six capteurs placés sur la surface de la paroi abdominale (voir figure 2.4, page 26) fournissent une information conjointe utilisable à la fois pour une validation croisée et pour une localisation des sources sonores à l'intérieur de l'abdomen. Le deuxième objectif est donc la construction d'une carte acoustique de l'activité abdominale à l'aide de la localisation.

4.1 Préliminaires : extraction des caractéristiques physiques

Les deux objectifs fixés ci-dessus nécessitent la prise en compte des caractéristiques physiques des sons abdominaux : d'une part, pour les distinguer des artéfacts, et d'autre part, pour évaluer l'effet de la propagation des sons à l'intérieur de l'abdomen et de localiser ainsi leur lieu d'origine.

[37] La version "minimale" de l'algorithme de débruitage est en phase de test dans la société Metravib SA pour la détection des bruits parasites dans les habitacles des voitures.

En conséquence, en amont de la re-segmentation et de la localisation, il est nécessaire d'extraire les caractéristiques physiques énumérées ci-dessous, qui semblent les plus pertinentes pour les experts médicaux et la bibliographie :

- la durée est la caractéristique la plus utilisée. Les valeurs données par la bibliographie varient de 20 ms. à 5 s. [48], mais sont généralement inférieures à 100 ms. [34, 84, 110, 116] ;
- l'énergie n'est pas adaptée pour décrire les événements individuellement, mais elle peut servir à l'analyse globale des successions de sons (dans le calcul de l'énergie par minute, par exemple). Cependant, les variations inter-individuelles risquent d'être importantes et de rendre l'information fournie inutilisable ;
- le spectre et l'argument de sa valeur maximale, la fréquence principale (*pitch*), se situent dans le domaine des basses-fréquences (inférieures à 1000 Hz.). Cette observation est confirmée par la bibliographie [15, 32, 74, 84, 116, 126], mais aussi par notre expérience (voir section 2.1) ;
- l'intensité acoustique, liée à l'amplitude de l'enveloppe de l'onde de pression sonore et donc à l'impression auditive (voir aussi la section 2.4.2, équation (2.7)), est relativement peu utilisée dans la littérature pour la caractérisation des sons abdominaux, pour les mêmes raisons que l'énergie. Cependant, elle constitue une caractéristique essentielle pour la localisation (voir section 4.3).

Ces caractéristiques permettent d'introduire des connaissances *a priori* dans la segmentation et elles servent à la localisation de la source sonore. Elles sont également à la base de la définition des indices d'activité. Ces indices, qui interviennent principalement dans l'analyse globale des phonoentérogrammes (la comparaison des signaux enregistrés sur les différentes voies, personnes, moments du cycle digestif, etc.), peuvent aussi être utilisés pour apprécier l'effet de la re-segmentation et de la localisation sur un même signal sonore abdominal.

Nous pouvons accéder aux caractéristiques physiques énumérées ci-dessus par (au moins) deux voies :

- la méthode classique, qui passe par la reconstruction du signal après le débruitage par ondelettes et sa segmentation à l'aide du vecteur binaire obtenu conformément à la section 3.6, consiste à calculer la durée, l'énergie, la fréquence principale et l'amplitude maximale de chaque événement à partir du signal reconstruit et segmenté. Outre le fait qu'elle implique la reconstruction du signal, elle réduit la description des sons abdominaux à ces caractéristiques physiques, utiles pour la segmentation et la localisation. De plus, elle peut être insuffisante pour une analyse plus détaillée et une classification ;
- une méthode directe, qui évite la reconstruction du signal et qui calcule les caractéristiques physiques à partir des coefficients d'ondelette de chaque événement. Vu que l'information fournie par la décomposition en ondelettes des sons abdominaux est complète, nous pouvons obtenir ces caractéristiques tout en conservant les coefficients d'ondelette pour les éventuelles étapes ultérieures de traitement.

Nous n'allons pas insister sur la méthode classique, car son application est immédiate. En revanche, nous allons montrer comment les caractéristiques physiques énumérées peuvent être extraites à partir des coefficients d'ondelette de chaque événement.

4.1.1 Ondelettes et caractéristiques physiques

La durée, l'énergie, la fréquence principale et l'intensité acoustique offrent une description limitée des sons abdominaux, peut-être insuffisamment adaptée à une classification performante, par exemple. Au contraire, les coefficients d'ondelette caractérisent précisément les événements et il apparaît donc souhaitable de garder cette information.

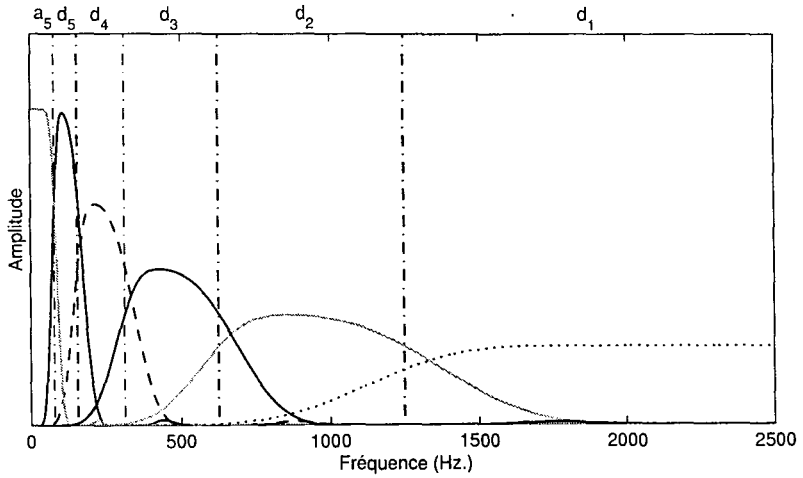


FIG. 4.1 – Spectres des ondelettes Daubechies no. 9 pour l'échelle de l'approximation et les 5 échelles des détails.

De plus, à ce stade du traitement, nous disposons déjà du vecteur contenant les coefficients de la décomposition en ondelettes du signal sonore abdominal débruité, ainsi que du vecteur binaire de segmentation, ce qui permet la construction d'un vecteur de coefficients d'ondelette pour chaque événement. A partir de ce vecteur, nous pouvons obtenir beaucoup d'informations sur les sons abdominaux, dont évidemment les caractéristiques physiques énumérées précédemment. Un premier pas consiste donc en l'identification des coefficients d'ondelette propres à chaque événement et leur extraction du vecteur des coefficients d'ondelette du signal débruité complet. Ce problème méthodologique revient en fait à la recherche des coefficients des ondelettes ayant contribué à la "construction" d'une zone particulière du signal.

Identification des coefficients d'ondelette d'un événement. La représentation bidimensionnelle classique d'une décomposition en ondelettes est présentée figure 4.2(a) : la décomposition est non-redondante, donc les "pavés" temps-fréquence (temps-échelle) sont adjacents et la couverture du plan complète. Cependant, cette figure est trompeuse, car elle nous laisse penser que deux lignes horizontales séparent une bande fréquentielle (une octave) et que deux lignes verticales délimitent un intervalle temporel. Ceci est vrai pour les bandes fréquentielles, sachant bien-sûr que les spectres des ondelettes des échelles voisines se superposent en partie (voir figure 4.1).

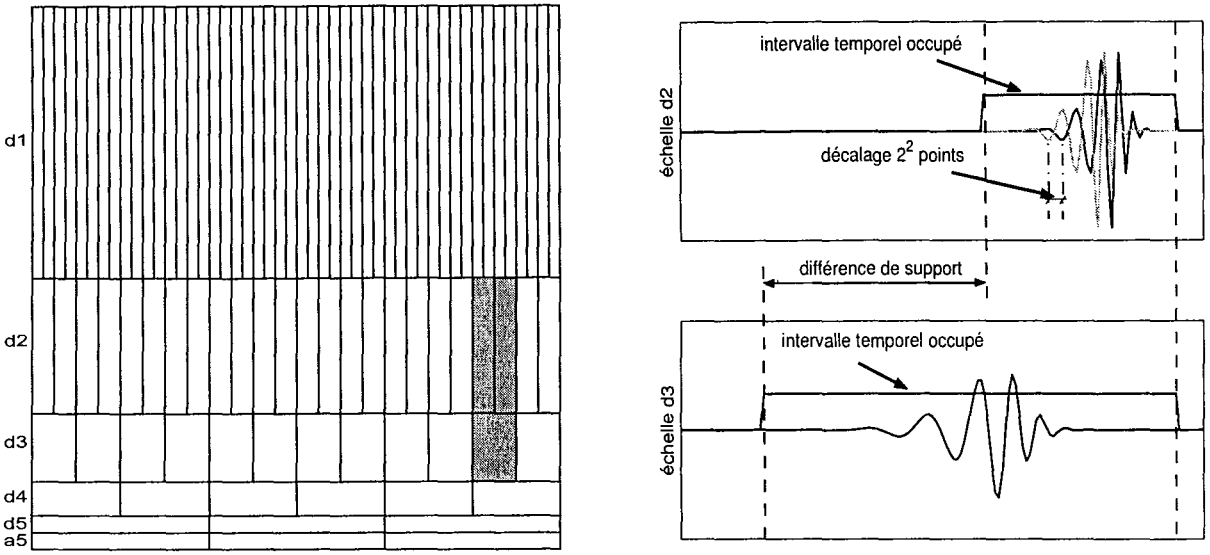
En revanche, la "colonne" constituée des pavés situés à la verticale d'un pavé de l'échelle d'approximation a_5 (un pavé sur l'échelle des détails d_5, deux sur d_4, jusqu'à 16 sur d_1) ne représente pas un intervalle temporel. Par exemple, les ondelettes correspondant aux 16 pavés de l'échelle haute-fréquence d_1 couvrent un intervalle temporel d'une longueur inférieure à l'intervalle correspondant aux 8 ondelettes de l'échelle d_2, et ainsi de suite.

L'explication est la suivante : pour un signal de longueur N, sur chaque échelle de détails[38] j de la décomposition il y a $N/2^j$ coefficients et le même nombre de pavés temps-fréquence (car chaque pavé correspond à un coefficient). Il y a donc $N/2^j$ ondelettes générées par ces coefficients. Cependant, la taille du support d'une ondelette est supérieure à 2^j points, puisqu'elle est donnée par son ordre et donc par la longueur de la réponse impulsionnelle du filtre de reconstruction (comme nous l'avons montré dans la section 3.6 dédiée à la segmentation par

[38]Dans la suite, nous allons écrire simplement j à la place de l'échelle de détails d_j, afin de simplifier la notation. L'échelle de l'approximation a_5 sera notée $M+1$ (où $M=5$ est la profondeur de la décomposition).

ondelettes, le nombre d'échantillons non nuls d'une ondelette de Daubechies à l'échelle j est de $T_j = (L-1)\cdot(2^j-1)+1$, où L est la longueur de la réponse impulsionnelle du filtre associé). Par exemple, les $N/2$ coefficients de l'échelle haute-fréquence d_1 génèrent $N/2$ ondelettes, décalées l'une par rapport à l'autre de 2 points mais ayant chacune une longueur égale à 18 points (échantillons), pour la base de Daubechies no. 9 que nous utilisons.

La règle générale est la suivante : les ondelettes générées par deux coefficients successifs situés à l'échelle j sont décalées d'un nombre 2^j de points, et ce décalage est inférieur à la taille de leur support. Par conséquent, à la différence des pavés temps-fréquence de la représentation classique (figure 4.2(a)), les ondelettes ne sont pas adjacentes mais superposées. D'un autre coté, on sait que la taille du support d'une ondelette à l'échelle $j+1$ est environ deux fois plus grande que celle d'une ondelette à l'échelle j. Comme deux ondelettes successives ne sont pas adjacentes, l'intervalle temporel qu'elles occupent est plus court que l'intervalle occupé par l'ondelette générée par le coefficient situé dans la même colonne mais sur l'échelle $j+1$ (voir figure 4.2(b)).



(a) pavés temps-fréquence

(b) échelles des détails et intervalles temporels

FIG. 4.2 – Relation entre les pavés temps-fréquence et les ondelettes correspondantes. Les deux pavés grisés de l'échelle d_2 (figure (a)) génèrent les deux ondelettes superposées (en haut de la figure (b)). Le pavé situé dans la même colonne mais dans la bande de fréquence inférieure d_3 génère une ondelette ayant un support de plus longue durée (figure (b), en bas) que les deux ondelettes superposées de d_2.

Les coefficients situés dans la même colonne du plan temps-échelle, mais à des échelles différentes, correspondent à des intervalles temporels de longueurs qui varient avec l'échelle. En particulier, en considérant les coefficients d'une seule colonne, plus la fréquence centrale de l'échelle augmente, plus l'intervalle temporel occupé par les ondelettes générées diminue. Réciproquement, si l'on veut retrouver les coefficients des ondelettes situées dans un même intervalle temporel, il est nécessaire de les chercher dans plusieurs colonnes. Nous présentons ensuite une méthode de calcul permettant la recherche des coefficients d'ondelette responsables d'un événement occupant un certain intervalle temporel, et donc la construction des représentations en ondelettes individuelles de chaque événement.

Considérons une ondelette à l'échelle $j+1$. La taille de son support temporel est $T_{j+1} = (L-1)\cdot(2^{j+1}-1)+1$, où L est la longueur de la réponse impulsionnelle du filtre de Daubechies

associé (18 points pour Daubechies no. 9). Considérons également une succession d'ondelettes à l'échelle j (de taille $T_j = (L-1)\cdot(2^j-1)+1$) ayant le même instant de fin (ou de début) que l'ondelette située à l'échelle plus basse fréquence $j+1$. Comme $T_j < T_{j+1}$ (voir aussi figure 4.2(b)), il est nécessaire de calculer combien d'ondelettes successives à l'échelle j occupent le même intervalle temporel que celui occupé par l'ondelette de l'échelle $j+1$, sachant que le décalage entre elles est de 2^j points. En notant n ce nombre, l'égalité d'intervalles temporels s'écrit :

$$T_{j+1} = T_j + n \cdot 2^j. \tag{4.1}$$

En remplaçant les expressions des tailles T_{j+1} et T_j, on obtient :

$$n = \frac{T_{j+1} - T_j}{2^j} = L - 1. \tag{4.2}$$

On peut donc avoir jusqu'à L ondelettes successives à l'échelle j responsables du même intervalle temporel qu'une seule ondelette de l'échelle $j+1$. Autrement dit, les deux coefficients (pavés) de l'échelle j situés "au dessus" du coefficient de l'échelle $j+1$ ne sont pas suffisants et il faut prendre encore $L-2$ coefficients, donc $\frac{L-2}{2}$ colonnes adjacentes.

Le paragraphe précédent a son importance dans la recherche des coefficients responsables d'un intervalle temporel donné du signal : par exemple, on ne peut pas considérer une "colonne" de pavés comme une fenêtre temporelle utilisable pour le calcul local de l'énergie ou pour la construction de l'enveloppe. Ce paragraphe permet également de proposer une méthode pour la construction d'un tableau de coefficients d'ondelette d'un événement, en considérant l'intervalle temporel (pour lequel on cherche les coefficients) donné par le vecteur binaire de segmentation.

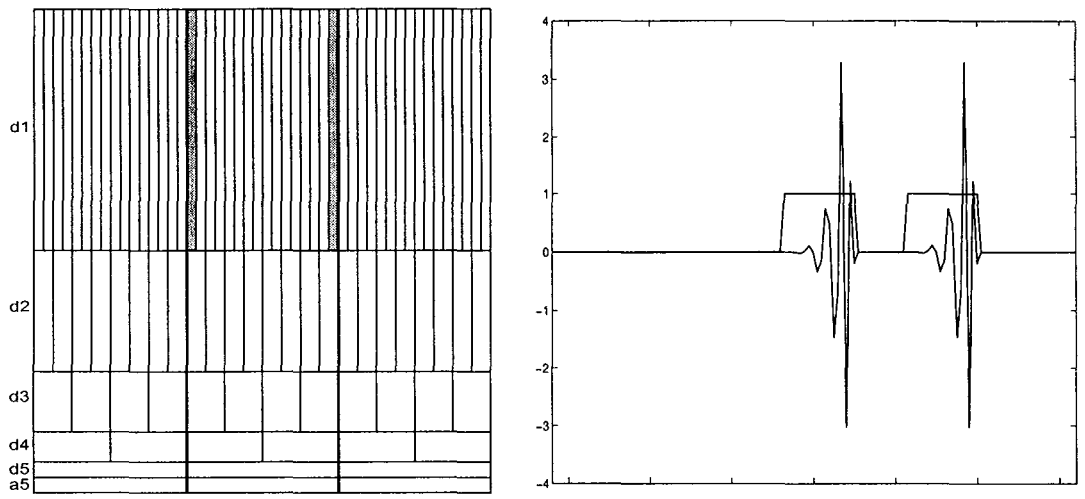
Avant de présenter l'algorithme d'extraction de coefficients d'un événement, il faut faire une remarque importante : les coefficients d'ondelettes d'une zone du signal (un événement) ne peuvent pas être extraits directement comme une suite de colonnes adjacentes de la représentation temps-échelle du signal entier. Par exemple, deux coefficients haute-fréquence situés dans la même colonne peuvent ne pas appartenir au même événement s'il n'y a pas un autre coefficient d'ondelette, plus basse fréquence, qui les relie (voir figure 4.3). Cependant, cette situation est détectée par l'algorithme de segmentation présenté dans la section 3.6.

On peut donc relier la méthode d'extraction des coefficients propres à un événement à la méthode de segmentation car, intuitivement, la première fait le chemin inverse de la deuxième : sur chaque échelle de la décomposition, il faut trouver les coefficients des ondelettes de l'événement étudié en prenant en compte les tailles et les positions temporelles de ces ondelettes.

L'algorithme de recherche des coefficients d'ondelette des événements passe donc par deux étapes :

1. Pour chaque événement, on identifie les colonnes de pavés temps-fréquence correspondant aux instants de début et de fin et on crée un vecteur de coefficients d'ondelette nuls $\{w_e\}$ ayant la dimension donnée par le nombre de coefficients situés dans ces colonnes.
2. Pour chaque échelle j, on recherche les coefficients d'ondelette responsables de l'intervalle temporel donné par le vecteur binaire de segmentation de l'événement, et on les affecte aux éléments correspondants de $\{w_e\}$.

A partir de ces vecteurs des coefficients d'ondelette de chaque événement, nous pouvons ensuite extraire les caractéristiques physiques nécessaires pour les étapes de re-segmentation et de localisation, voire pour la classification des sons abdominaux.



(a) pavés temps-fréquence (b) intervalles temporels

FIG. 4.3 – Disposition des pavés temps-fréquence dans le cas d'une représentation par ondelettes (a). Le signal segmenté, généré par les deux pavés grisés de la même colonne, est présenté dans (b).

Durée. Le calcul des instants de début et de fin et donc de la durée d'un son abdominal à partir du vecteur binaire de segmentation obtenu dans la section 3.6 est évidente. Pour un événement donné $e(k)$, nous avons accès aux instants de début et de fin et donc à sa taille N_e. La durée en secondes sera obtenue comme $D_e = N_e/f_s$, où $f_s = 5000$ Hz. désigne la fréquence d'échantillonnage.

Énergie. Comme nous l'avons vu dans la section consacrée aux bases orthogonales d'ondelettes (section 3.1.2), un signal et donc un événement $e(k)$, $k = 1 \ldots N_e$, peut s'écrire :

$$e(k) = \sum_{j,p} w_e^{j,p} \psi_{j,p}(k) + \sum_{p} w_e^{M+1,p} \phi_{M+1,p}(k), \tag{4.3}$$

où $\psi_{j,p}$ (respectivement $\phi_{M+1,p}$) désignent les ondelettes (les fonctions d'échelle) et $w_e^{j,p}$ les coefficients d'ondelette de l'échelle j et position temporelle p.

L'énergie, conservée par projection orthogonale, est égale à :

$$E_e = \sum_{k} |e(k)|^2 = \sum_{j,p} \left(w_e^{j,p}\right)^2. \tag{4.4}$$

Spectre et fréquence principale. La disposition des pavés temps-fréquence (figure 4.2(a)) et l'algorithme de la décomposition en ondelettes orthogonales par filtrage nous permettent de lier l'énergie d'une bande fréquentielle aux coefficients d'ondelette de l'échelle correspondante :

$$E_e^j = \sum_{p} \left(w_e^{j,p}\right)^2, \quad \text{avec} \quad j = 1 \ldots M+1. \tag{4.5}$$

En prenant en compte la largeur de bande spécifique à la résolution fréquentielle variable offerte par la décomposition en ondelettes, nous pouvons construire une approximation de l'enveloppe spectrale (figure 4.4) qui peut servir à la détection approximative de la fréquence principale.

Cependant, vu la faible résolution de cette approximation ($M+1$ valeurs entières possibles), le maximum détecté n'est pas vraiment discriminant, surtout que, selon notre expérience (section 2.1), une grande partie de l'énergie des sons abdominaux se concentre dans deux bandes fréquentielles : $j = 4$, donc ≈[157-312] Hz. et $j = 3$, ≈[312-625] Hz.

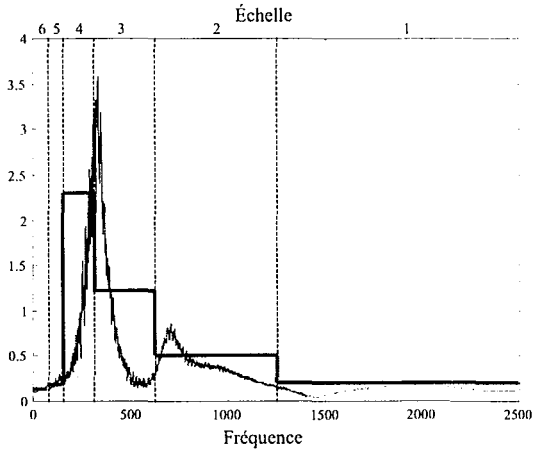


FIG. 4.4 – Spectre et enveloppe spectrale d'un événement

Une autre solution est la reconstruction exacte du spectre de l'événement étudié. Pour un signal quelconque, sa transformée de Fourier est égale à la somme des transformées de Fourier des ondelettes qui le composent. Pour un événement $e(k)$, cette observation s'écrit :

$$\hat{e}(f) = \sum_{j,p} w_e^{j,p} \hat{\psi}_{j,p}(f) + \sum_{p} w_e^{M+1,p} \hat{\phi}_{M+1,p}(f), \tag{4.6}$$

où $\hat{\psi}_{j,p}(f)$ et $\hat{\phi}_{M+1,p}(f)$ sont respectivement les transformées Fourier des ondelettes et des fonctions d'échelle.

La fréquence principale est alors donnée par la position du maximum du module de la transformée de Fourier (éq. 4.6) :

$$f_e = \arg\max_f |\hat{e}(f)| . \tag{4.7}$$

Amplitude et intensité acoustique. L'énergie d'un son sert à la construction des indices d'activité abdominale (voir section 4.1.2) et au calcul de sa puissance, mais elle n'est pas la caractéristique la mieux adaptée pour la classification : un médecin classe plutôt un son en fonction de l'impression auditive, donc de son intensité acoustique (éq. 2.7). De plus, comme nous allons le montrer dans la section 4.3, c'est l'intensité acoustique (directement liée à l'amplitude de l'onde de pression) qui varie par propagation et qui est utilisée dans la localisation des sources des sons abdominaux.

Pour simplifier, un événement $e(k)$ ayant une intensité acoustique variable en temps, nous avons choisi de le caractériser par l'intensité acoustique maximale, liée à l'amplitude maximale de son enveloppe. Une méthode possible pour obtenir l'enveloppe passe par le calcul de l'énergie sur des fenêtres glissantes. Comme pour le spectre, nous proposons ensuite une approche basée sur les propriétés de la décomposition en ondelettes.

Le calcul de l'enveloppe d'un événement est lié à sa segmentation : pour chaque échelle j, on peut considérer une fenêtre glissante ayant la taille du support temporel des ondelettes de l'échelle associée. Si le pas de glissement est 2^j, cela revient à considérer le vecteur de segmentation de

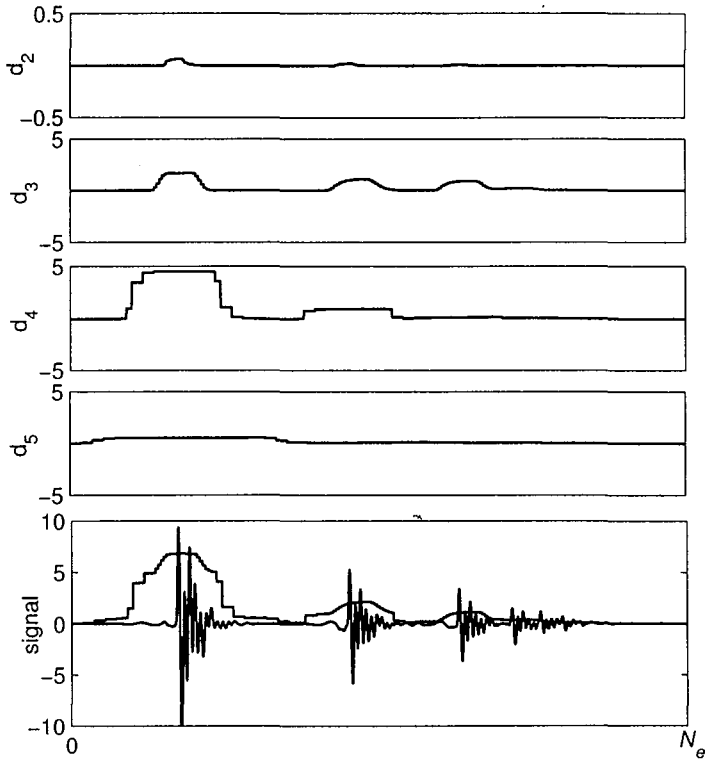


FIG. 4.5 – Enveloppes des échelles de détails (pour les échelles non-nulles) et de l'événement reconstruit. L'enveloppe de l'échelle d_2 est agrandie 10 fois.

chaque ondelette. Dans ce cas, l'énergie de chaque fenêtre est donnée directement par le carré du coefficient d'ondelette correspondant. Reste le problème de la résolution : comme pour l'enveloppe spectrale, elle varie en fonction de l'échelle de la décomposition. En pratique, l'algorithme que nous proposons reprend la méthode de segmentation, avec une différence importante : il ne s'agit plus d'un OU logique entre les vecteurs binaires de segmentation de chaque échelle (voir section 3.6), mais d'une somme entre les vecteurs de segmentation modifiés pour prendre en compte l'énergie des ondelettes (figure 4.5).

Les étapes de l'algorithme du calcul de l'enveloppe sont :

1. Initialiser les vecteurs d'enveloppes des échelles à 0.
2. Pour chaque échelle j :
   (a) pour chaque coefficient non nul $w_e^{j,p}$, calculer les instants de début u et de fin d de l'ondelette associée et son vecteur binaire de segmentation ;
   (b) modifier l'amplitude de ce vecteur binaire de sorte que son énergie soit égale à celle de l'ondelette associée (multiplication par $w_e^{j,p}$ et normalisation par la taille du support $d - u + 1$) ;
   (c) additionner les vecteurs calculés précédemment afin d'obtenir l'enveloppe de l'échelle courante (grâce à l'orthogonalité des ondelettes, l'énergie de l'enveloppe est égale à celle de l'échelle).
3. Additionner les enveloppes des échelles pour obtenir l'enveloppe finale de l'événement (d'énergie égale à celle de l'événement).

Comme dans le cas de l'enveloppe spectrale, nous avons aussi un problème de résolution. Cependant, la relation échelle-résolution est inverse : plus l'énergie de l'événement est située dans les basses fréquences, moins sa localisation temporelle est précise. Qui plus est, la méthode proposée redistribue l'énergie de chaque ondelette sur tout le support donné par le vecteur de segmentation[39]. Autrement dit, pour chaque ondelette, l'enveloppe est un signal rectangulaire avec une amplitude donnée par son énergie moyenne, et l'enveloppe finale de l'événement est la somme de ces signaux rectangulaires. Ceci n'est pas une méthode précise de calcul de l'enveloppe et donc de l'intensité acoustique d'un son abdominal.

Comme pour la fréquence maximale (qui peut être obtenue à partir du spectre du signal $e(k)$ reconstruit), l'amplitude maximale peut être obtenue après reconstruction (voire après une transformée de Hilbert appliquée au signal reconstruit pour obtenir son enveloppe).

Par conséquent, la meilleure solution pour le calcul de l'intensité acoustique est, comme dans le cas de la fréquence maximale, la reconstruction exacte de l'événement étudié (éq. 4.3). L'intensité acoustique s'écrit donc :

$$I_e = \max_k (e(k))^2. \tag{4.8}$$

4.1.2 Indices d'activité

Quelle que soit la méthode d'extraction des caractéristiques physiques des événements, leur prise en compte sur de longues durées d'enregistrement permet d'apprécier l'activité globale de l'appareil digestif. Ainsi, nous pouvons définir plusieurs indices d'activité liés à la motricité abdominale :

- le nombre d'événements par unité de temps N_m ;
- l'énergie totale par unité de temps $E_m = \sum E_e$;
- la durée totale des événements rapportée à l'unité de temps, $D_m = \frac{1}{T_m} \sum D_e \cdot 100$.

Cette liste d'indices n'est bien sûr pas exhaustive. Qui plus est, ces indices ne sont peut-être pas les meilleurs pour une description à visée physiologique – ils ne prennent pas en compte et ne détectent pas la rythmicité des contractions gastro-intestinales, par exemple. Cependant, nous les considérons suffisants pour une première approche descriptive des phonoentérogrammes. Leur objectif principal est, dans ce cas, non pas de faciliter l'interprétation physiologique mais de fournir des critères quantitatifs pour l'évaluation des signaux durant les étapes de traitement suivantes, la re-segmentation et la localisation.

A part les indices principaux, qui caractérisent quantitativement l'activité sonore abdominale, nous pouvons définir des indices secondaires, moins utilisés dans la bibliographie mais peut-être informatifs pour la description de la physiologie gastro-intestinale[40] :

- l'énergie moyenne des événements apparaissant par unité de temps $E_\mu = E_m / N_m$;
- la durée moyenne des événements $D_\mu = \frac{1}{N_m} \sum D_e$;
- la puissance moyenne des événements $P_\mu = \frac{1}{N_m} \sum \frac{E_e}{D_e}$;
- la fréquence principale moyenne des événements $f_\mu = \frac{1}{N_m} \sum f_e$;
- l'intensité acoustique moyenne des événements $I_\mu = \frac{1}{N_m} \sum I_e$;
- la durée moyenne des silences séparant les événements $D_{s,\mu}$.

Nous n'allons pas les utiliser pour évaluer les étapes de re-segmentation et de localisation, mais ils auront leur rôle dans le chapitre dédié à l'analyse des signaux (chapitre 5).

[39] En réalité, l'énergie d'une ondelette n'est pas répartie uniformément sur son support, et sa puissance est concentrée dans un intervalle temporel inférieur à son support et généralement pas centré.

[40] L'unité de temps T_m couramment considéré dans la littérature est la minute.

3. **Élimination des sons de basse-fréquence.** Cette étape ne propose pas un filtrage passe-haut symétrique à la première étape décrite ci-dessus, équivalent à l'annulation des coefficients d'ondelette situés sur l'échelle basse-fréquence de l'approximation a_5. Si presque aucun son abdominal n'a de composante spectrale supérieure à 1250 Hz., ce n'est pas le cas des composantes basse-fréquence (inférieures à 80 Hz.), qui sont présentes dans la plupart des sons physiologiques, sans pour autant être très énergétiques (car les fréquences principales des sons abdominaux sont supérieures à 80, voire à 100 Hz.). Nous considérons qu'un filtrage passe-haut risque d'éliminer une information qui peut être importante. En effet, cette information est plus de nature "structurelle" que sonore : les ondelettes de grande taille générées par les coefficients de l'échelle de l'approximation ont un rôle important dans la concaténation des micro-événements d'ailleurs inséparables à l'écoute et donc dans la segmentation. Par conséquent, nous ne pouvons pas annuler directement les coefficients sur l'échelle de l'approximation sans perdre cette information. En revanche, nous proposons l'utilisation de l'enveloppe spectrale calculée à l'aide de l'expression (4.5) pour détecter si l'énergie d'un événement est située plutôt dans les basses fréquences. Ainsi, aux critères d'élimination d'un artéfact se rajoute le rapport entre l'énergie des basses fréquences et l'énergie totale : si plus de la moitié de l'énergie est située en dessous de 80 Hz., l'événement est considéré comme un artéfact[42].

4. **Validation croisée.** La dernière étape d'élimination des artéfacts consiste en la vérification d'une condition de propagation. L'hypothèse de base est que les sons abdominaux (du moins ceux qui ont une intensité acoustique suffisamment importante pour les rendre informatifs) se propagent à l'intérieur de l'abdomen et sont par conséquent enregistrés par au moins deux capteurs. Autrement dit, pour être validé, le même événement doit être présent simultanément sur au moins deux voies ; sinon, il est considéré comme un artéfact. Reste à définir ce qu'on entend par "le même" et par "simultanément". Une observation importante mérite d'être faite : la simultanéité temporelle, qu'on peut définir comme une superposition plus ou moins importante ou encore comme le résultat d'un OU logique entre les vecteurs binaires de segmentation de deux événements, n'est pas *a priori* la preuve de l'existence d'un son abdominal propagé. En effet, on peut imaginer une superposition temporelle de deux sons d'origine différente mais concomitants. Dans ce cas, on devrait calculer par exemple un coefficient d'inter-corrélation entre les deux événements en question (ou entre leurs coefficients d'ondelette) et décider ensuite s'il s'agit d'une propagation. Cependant, les signaux sonores abdominaux sont supposés creux et la probabilité d'apparition d'un son abdominal est considérée comme faible[43]. De plus, si on respecte des conditions d'enregistrement correctes, on peut supposer que la probabilité d'apparition d'un artéfact est encore plus faible que celle d'un vrai événement. Par conséquent, nous pouvons faire l'hypothèse qu'un événement qui n'est pas simultané en temps à un autre est un artéfact et donc qu'il peut être éliminé. Cette hypothèse ne nous assure pas de l'élimination de tous les artéfacts, car il est toujours possible qu'un son abdominal soit simultané à un artéfact. Cependant, la probabilité conjointe d'apparition est extrêmement faible.

   Cette dernière étape de l'algorithme d'élimination des artéfacts consiste donc en l'annulation de tous les événements présents sur une seule voie, c'est-à-dire dont le support temporel ne se superpose au support d'aucun autre événement sur aucune autre voie.

[42] Une autre solution possible est de prendre en compte la fréquence principale et d'éliminer les événements qui ont cette fréquence inférieure à 80 ou à 100 Hz.

[43] Cette affirmation est à prendre avec précaution, car la durée totale des événements après le débruitage peut atteindre plus de 20% de la durée du signal (voir figure 3.27). Cependant, vu que cette étape est la dernière de l'algorithme de re-segmentation, la durée totale est déjà réduite par les traitements précédents.

4.2.2 Résultats, exemples

Le but de cette section est d'évaluer l'effet des quatre étapes de re-segmentation décrites précédemment. Pour quantifier cet effet, nous utilisons les indices d'activité principaux introduits dans la section 4.1.2 et nous appliquons les étapes de l'algorithme de re-segmentation l'une après l'autre, en les évaluant individuellement. Nous illustrons les effets de la re-segmentation sur un exemple : un phonoentérogramme à six voies, d'une durée d'environ 168 minutes. Nous évaluons les valeurs moyennes des indices principaux N_m, D_m et E_m (voir section 4.1.2) pour les six voies, sur les 168 minutes d'enregistrement. Ces évaluations sont effectuées avant toute élimination (donc immédiatement après le débruitage) et après chaque étape d'élimination des artéfacts. Les résultats détaillés par voie sont présentés dans les tableaux 4.1 à 4.5. Les valeurs moyennes pour les six voies confondues sont présentées sur les figures 4.6, 4.7(a) et 4.7(b).

Un premier constat, attendu, est que pour les trois critères les valeurs décroissent à chaque étape. Cependant, les deux premiers indices ont une évolution très semblable, alors que la décroissance de l'énergie totale est beaucoup moins importante : si le nombre d'événements par minute et la durée totale des événements par minute se réduisent après re-segmentation à environ un tiers de leurs valeurs initiales, l'énergie totale perd moins de 2,5%. Autrement dit, les événements éliminés sont généralement de faible énergie (une grande partie de ces événements semble donc constituée de fausses alarmes de l'algorithme de débruitage).

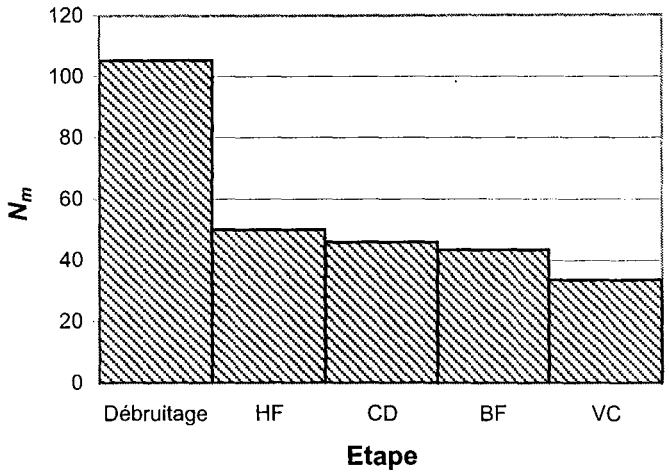
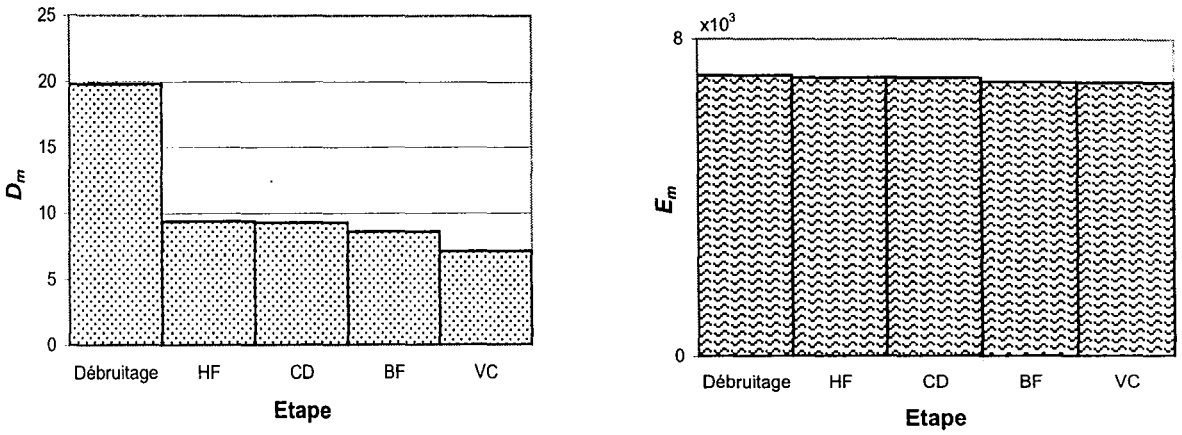


FIG. 4.6 – Évolution de l'indice N_m – nombre d'événements par minute. Les étapes de l'algorithme d'élimination d'artéfacts sont marquées en abscisse : HF = annulation des coefficients de haute-fréquence, CD = élimination des événements de courte durée, BF = élimination des événements de basse-fréquence, VC = validation croisée. L'ordonnée représente la valeur de l'indice après l'étape respective.

Le saut le plus important en ce qui concerne la durée D_m (et le nombre d'événements N_m) est dû à l'annulation des ondelettes haute-fréquence. Il s'agit soit des fausses-alarmes issues de l'algorithme de débruitage, soit des micro-frottements de la tête de stéthoscope contre la peau.

L'étape suivante (l'élimination des événements de courte durée) laisse les signaux quasiment inchangés. Elle garde un éventuel intérêt pour l'élimination des artéfacts de l'échelle de fréquence immédiatement inférieure ([625–1250] Hz.), qui ont encore un support plus court que 20 ms.

L'élimination des artéfacts basse-fréquence ne change presque pas les deux premiers indices d'activité, car ces événements sont *a priori* rares si l'enregistrement est de bonne qualité. Cependant, le saut énergétique est important (voir les tableaux 4.3 et 4.4), voire le plus important de toutes les étapes de re-segmentation. Comme nous l'avons précisé précédemment, il s'agit



(a) durée totale par minute

(b) énergie totale par minute

FIG. 4.7 – Évolution des indices D_m (durée totale par minute) et E_m (énergie totale par minute). Les ordonnées représentent les valeurs respectives des indices après les différentes étapes d'élimination d'artéfacts.

souvent de sons de forte intensité (voir l'exemple présenté figure 4.8) provoqués par des chocs ou des frottements directs sur la tête de stéthoscope.

Enfin, l'étape de validation croisée affecte de nouveau très peu l'énergie totale E_m mais davantage les deux autres indices d'activité (N_m et D_m), ce qui est bien-sûr attendu : seuls les artéfacts ou les événements de très faible intensité ne se propagent pas dans l'abdomen et ne sont pas détectés par plusieurs capteurs.

Une étude détaillée des indices d'activité par voie nous dévoile des aspects intéressants : après la première étape de re-segmentation (l'élimination des ondelettes haute-fréquence), la variation du nombre d'événements par minute est beaucoup plus importante pour la voie no. 1 que pour les autres. L'explication, pour cet exemple particulier de phonoentérogramme, est fournie par l'écoute du signal enregistré par le premier capteur : il est fortement parasité par des bruits courts et aigus dus à un défaut de contact entre la membrane du stéthoscope et la peau (pression d'appui trop faible). Comme ce capteur est proche de la cage thoracique, le signal qu'il fournit est le plus affecté par les mouvements respiratoires, qui arrivent à provoquer des micro-frottements quand la pression d'appui est trop faible.

A part ce cas particulier, l'évolution des indices d'activité semble similaire pour les six voies. Comme les trois premières étapes de re-segmentation traitent les six enregistrements indépendamment l'un de l'autre, la situation paraît logique. La dernière étape (la validation croisée) n'affecte pas d'une manière évidente une des voies en particulier. En effet, les événements éliminés par validation croisée peuvent être de véritables sons abdominaux, vu que leurs caractéristiques physiques n'ont pas conduit à leur élimination dans une étape antérieure. Sous cet angle, la dernière étape de re-segmentation peut être vue comme une élimination des sons de très faible intensité acoustique. Elle revient à l'introduction d'une condition supplémentaire d'intensité minimale, qui se rajoute aux conditions sur le contenu fréquentiel et sur la durée imposées aux étapes précédentes de l'algorithme d'élimination des artéfacts.

Indice	Voie						Moyenne
	1	2	3	4	5	6	
N_m	238,2	69,1	36,9	91,0	95,0	101,1	105,3
D_m	18,5%	16,3%	8,2%	31,9%	25,3%	18,6%	19,8%
E_m	9498	5106	1870	5226	8795	12 054	7091

TAB. 4.1 – Exemple de valeurs des principaux indices d'activité après débruitage (valeurs moyennes sur 168 minutes d'enregistrement).

Indice	Voie						Moyenne
	1	2	3	4	5	6	
N_m	38,8	36,1	33,2	70,0	66,9	54,8	50,0
D_m	7,4%	7,9%	4,1%	15,8%	12,5%	9,0%	9,4%
E_m	9366	5048	1859	5195	8756	11 957	7030

TAB. 4.2 – Exemple de valeurs des principaux indices d'activité après la première étape de re-segmentation : élimination des ondelettes haute-fréquence (valeurs moyennes sur 168 minutes d'enregistrement).

Indice	Voie						Moyenne
	1	2	3	4	5	6	
N_m	34,6	33,5	32,5	67,3	57,6	49,9	45,9
D_m	7,3%	7,9%	4,0%	15,7%	12,3%	8,9%	9,3%
E_m	9364	5047	1859	5195	8755	11 957	7030

TAB. 4.3 – Exemple de valeurs des principaux indices d'activité après l'élimination des événements de courte durée (valeurs moyennes sur 168 minutes d'enregistrement).

Indice	Voie						Moyenne
	1	2	3	4	5	6	
N_m	32,4	30,8	32,3	62,6	56,4	45,8	43,4
D_m	6,4%	7,2%	4,0%	14,4%	11,8%	7,8%	8,6%
E_m	9222	5043	1859	4961	8597	11 932	6936

TAB. 4.4 – Exemple de valeurs des principaux indices d'activité après l'élimination des événements de basse fréquence (valeurs moyennes sur 168 minutes d'enregistrement).

Indice	Voie						Moyenne
	1	2	3	4	5	6	
N_m	26,8	27,4	25,7	41,5	44,2	35,8	33,6
D_m	5,5%	6,6%	3,6%	10,3%	10,0%	6,7%	7,1%
E_m	9219	5040	1855	4934	8572	11 918	6923

TAB. 4.5 – Exemple de valeurs des principaux indices d'activité après validation croisée (valeurs moyennes sur 168 minutes d'enregistrement).

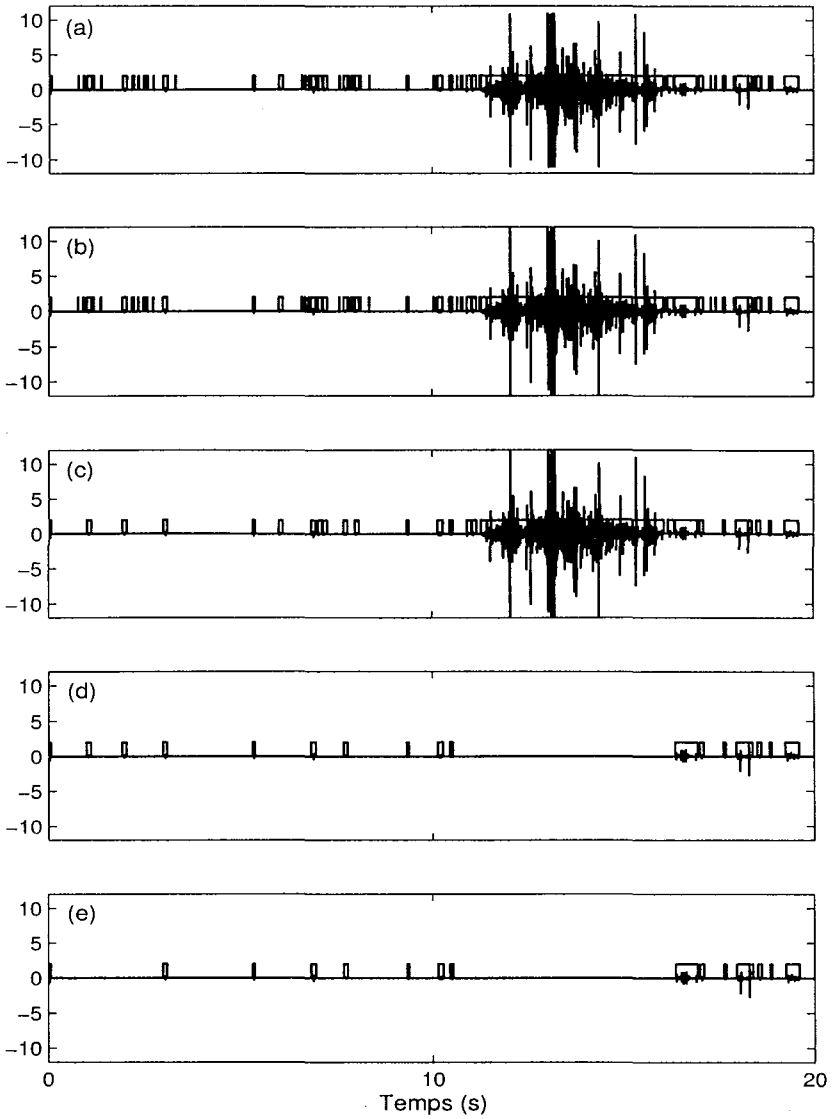


FIG. 4.8 – Exemple de re-segmentation (les valeurs des indices d'activité principaux sont données entre parenthèses) : (a) signal après débruitage (N_m = 68, D_m = 42,1%, E_m = 67849,4 V^2) ; (b) après élimination des hautes-fréquences (N_m = 65, D_m = 41,9%, E_m = 66630,4 V^2) ; (c) après élimination des événements de courte durée (N_m = 25, D_m = 41,1%, E_m = 66630,3 V^2) ; (d) après élimination des événements basse-fréquence (N_m = 17, D_m = 13,1%, E_m = 272,6 V^2) ; (e) après validation croisée (N_m = 15, D_m = 11,9%, E_m = 270,5 V^2). Pour cet exemple, l'élimination des ondelettes haute-fréquence n'apporte pas d'amélioration notable. En revanche, le signal se modifie par l'élimination des événements de courte durée, et par validation croisée. L'effet le plus important est l'élimination d'un artéfact de forte énergie dû à un mouvement du patient, grâce à l'étape de détection des événements basse-fréquence.

4.3 Localisation

Relativement peu de travaux dans le domaine de la phonoentérographie prennent en compte l'aspect localisé des phénomènes digestifs. Les méthodes de localisation ne sont pas précises : en général, l'abdomen est partagé en plusieurs régions d'intérêt, qui doivent être auscultées successivement et pendant au moins deux minutes chacune (voir par exemple la description des techniques d'auscultation dans [2] et [49]).

Dans certaines études, les auteurs se focalisent sur deux sites principaux, notamment la région épigastrique (partie inférieure de l'estomac/duodénum) et la région péri-ombilicale (voire souvent infra-ombilicale) [34, 116, 119, 121] et ils signalent des différences d'activité quantifiables par différents indices (nombre de sons par minute, énergie moyenne, durée totale des sons, fréquence moyenne). D'autres équipes réalisent des enregistrements en trois ou quatre sites (épi-gastrique, quadrant inférieur gauche, quadrant inférieur droit ...), en détectant toujours des différences d'activité [15, 76, 99]. Cette démarche prouve l'intérêt de l'information spatiale dans l'investigation médicale, d'autant plus que certains auteurs relient les différents types de manifestations sonores à des organes précis de la cavité abdominale [18, 41].

Cependant, les sons abdominaux se propagent à l'intérieur de l'abdomen et donc les enregistrements réalisés en un site contiennent des sons générés par des organes situés aux différents endroits de la cavité abdominale. Il est intéressant de localiser les sources des sons, car on peut supposer que chaque son est généré par un phénomène digestif lié à un mouvement et donc à une contraction d'un organe situé à un endroit précis de la cavité abdominale. La prise en compte de l'aspect multi-voies des enregistrements, déjà initiée par la validation croisée décrite précédemment, peut améliorer sensiblement le traitement des signaux sonores abdominaux par localisation. Nous pouvons ainsi extraire une information locale, qui s'ajoute aux indices d'activité pour permettre la construction d'une carte abdominale dynamique, évoluant dans le temps. Cet objectif est basé sur l'hypothèse de la rareté des sons abdominaux : comme le signal est creux, on peut faire l'hypothèse que la probabilité d'apparition concomitante de deux ou plusieurs événements à deux endroits différents est faible, et donc que tout événement présent simultanément sur deux voies est dû à un seul phénomène physique (un seul son), dont la source nécessite d'être localisée.

Comme pour la validation croisée, la simultanéité de deux événements est basée sur leur superposition temporelle. Les événements non-propagés étant éliminés après l'algorithme de re-segmentation, tous les sons restants seront présents sur au moins deux voies.

4.3.1 Méthodes de localisation

Nous présentons dans cette section, brièvement et d'une manière simplifiée, les principales méthodes de localisation. Le milieu de propagation considéré est supposé isotrope et invariant dans le temps[44].

En général, la localisation d'une source quelconque consiste en l'estimation des distances entre la position (inconnue) de cette source et les positions (connues) des capteurs[45]. Dans un espace cartésien tridimensionnel, la distance d_i entre la source et le capteur i s'écrit :

$$d_i = \sqrt{(x_i - x_s)^2 + (y_i - y_s)^2 + (z_i - z_s)^2}, \tag{4.9}$$

[44]En fait, les problèmes posés par les sons abdominaux ont une spécificité qui rend difficile l'application des méthodes classiques, comme nous allons le montrer.

[45]Les méthodes d'estimation de la direction ou de l'angle d'arrivée d'un signal (*angle of arrival - AOA* ou *direction of arrival - DOA*) impliquent l'utilisation de capteurs directionnels et ne seront pas abordées ici.

où (x_i, y_i, z_i) et (x_s, y_s, z_s) sont respectivement les coordonnées du capteur i et de la source. Comme les positions des capteurs sont connues, l'estimation de la distance d_i revient au calcul des coordonnées (x_s, y_s, z_s) de la source.

Pour cette estimation, on peut recourir à plusieurs indices [24, 66, 92, 103] :

1. Le temps d'arrivée (*time of arrival - TOA*), qui consiste en l'estimation de la distance entre la source et le capteur en fonction du temps nécessaire au signal pour la parcourir. Cette étape place la source à l'intersection des sphères centrées autour des capteurs avec des rayons estimés par les calculs de distance précédents. Évidemment, comme nous ne connaissons pas le moment exact d'émission du son étudié, nous ne pouvons pas appliquer cette méthode dans notre cas.
2. La différence des temps d'arrivée (*time delay of arrival - TDOA*), qui estime la différence entre deux distances d_i et d_j en fonction de la différence entre les temps d'arrivée du signal aux capteurs. Notons Δt_i le temps (inconnu) entre l'émission du son et son arrivée au capteur i. La distance d_i sera écrite :

   $$d_i = c \cdot \Delta t_i,$$

   où c désigne la célérité (la vitesse moyenne de propagation) du son dans le milieu considéré. Il est immédiat que

   $$\begin{aligned} d_i - d_j &= \sqrt{(x_i - x_s)^2 + (y_i - y_s)^2 + (z_i - z_s)^2} - \sqrt{(x_j - x_s)^2 + (y_j - y_s)^2 + (z_j - z_s)^2} \\ &= c \cdot (\Delta t_i - \Delta t_j) = c \cdot \Delta t_{ij}, \end{aligned} \tag{4.10}$$

   où Δt_{ij} désigne la différence de temps d'arrivée mesurée. Par conséquent, comme cet intervalle temporel est connu, la différence entre les deux distances est connue aussi et la position de la source sonore est obtenue comme l'intersection des surfaces hyperboliques (4.10).

   Un problème important soulevé par cette approche est la méthode de calcul de la différence de temps d'arrivée. En effet, la *TDOA* peut être calculée comme l'intervalle entre les instants de début des sons enregistrés sur chacune des voies, ou encore comme la position du pic de leur fonction d'inter-corrélation.
3. Le rapport d'énergies ou d'intensités acoustiques, qui renonce aux mesures de retard temporel pour estimer les distances, en introduisant à la place des hypothèses sur l'atténuation énergétique par propagation. Le modèle classique de propagation dans un milieu isotrope non-absorbant [81] s'écrit :

   $$I_i = \frac{P_s}{4\pi d_i^2}, \tag{4.11}$$

   où I_i est l'intensité acoustique du son enregistré par le capteur i, P_s la puissance de la source sonore, d_i la distance entre la source et le capteur (4.9)[46]. En suivant un développement analogue à celui de Li [66], la puissance inconnue de la source peut être calculée à partir de chaque signal d'intensité I_i enregistré par le capteur i. Pour une paire de capteurs (i, j), l'expression obtenue :

   $$4\pi d_i^2 I_i = 4\pi d_j^2 I_j = P_s, \tag{4.12}$$

   représente l'intersection de deux sphères de rayons inversement proportionnels aux racines carrées des intensités des sons enregistrés.

[46] En acoustique, cette équation est souvent écrite directement pour les amplitudes des ondes de pression, qui varient de façon inversement proportionnelle à la distance [104].

La position de la source est retrouvée, comme pour la méthode des temps d'arrivée (TOA), à l'intersection des sphères, c'est-à-dire en résolvant le système d'équations non-linéaires d'inconnues (x_s, y_s, z_s), obtenu pour toutes les paires (i, j) en remplaçant (4.9) dans (4.12) :

$$I_i \left((x_i - x_s)^2 + (y_i - y_s)^2 + (z_i - z_s)^2\right) - I_j \left((x_j - x_s)^2 + (y_j - y_s)^2 + (z_j - z_s)^2\right) = 0. \tag{4.13}$$

A la différence de TOA, qui demande $N + 1$ capteurs pour un espace de dimension N, cette méthode nécessite $N + 2$ capteurs, à cause de la nouvelle inconnue P_s.

A notre connaissance, dans le domaine de la phonoentérographie seulement deux publications abordent la localisation des sons :

- une démarche empirique mais apparentée à la méthode des différences de temps d'arrivée ($TDOA$) a été suggérée par Garner et Ehrenreich [48], qui utilisent le décalage entre les instants de début de deux événements segmentés et la vitesse du son dans l'air pour avancer l'hypothèse d'une propagation des sons en fonction de l'anatomie abdominale, le long des intestins. Cependant, leur approche n'est pas systématique et ils présentent un seul exemple de son propagé selon cette hypothèse. En pratique, comme le contenu des intestins varie en temps, l'hypothèse de Garner et Ehrenreich ne peut pas être vérifiée et elle n'est pas applicable pour la localisation des sources.
  Sous l'hypothèse simplifiée de propagation dans un milieu isotrope, la méthode $TDOA$ devrait rester intéressante. Cependant, la vitesse de propagation des ondes sonores dans les tissus humains mous est proche de la vitesse du son dans l'eau, c'est-à-dire 1500 m/s [52]. A cette vitesse, les distances inférieures à 1/3 m sont parcourues en moins de 0,2 ms. Autrement dit, le pas d'échantillonnage de notre système d'acquisition (fréquence d'échantillonnage de 5000 Hz.) ne nous permet pas la détection des différences temporelles nécessaires pour l'application de cette méthode ;
- la méthode des rapports d'intensité acoustique, introduite dans le domaine de la phonoentérographie par Craine *et al.* [32], est la seule utilisée systématiquement pour la localisation des sons abdominaux. Le modèle utilisé par les auteurs de [32] suppose trois capteurs autour d'une source ponctuelle, dans un environnement isotrope bidimensionnel. La solution trouvée est unique, car la contrainte implicite imposée par les auteurs fixe la position recherchée à l'intérieur du triangle équilatéral formé par les capteurs (les stéthoscopes 1, 3 et 5 dans la figure 2.4).

4.3.2 Améliorations

L'augmentation du nombre de capteurs (Fig. 2.4) ouvre des perspectives intéressantes. Premièrement, nous pouvons aborder la localisation en profondeur des sources sonores en rajoutant une troisième dimension à l'espace proposé dans [32]. Ensuite, par la redondance introduite, elle permet de tester la validité de la méthode de localisation et implicitement du modèle de propagation choisi (de tester la robustesse de l'estimée de la position). Les résultats de ces tests permettent d'apporter d'éventuelles modifications et de choisir finalement la méthode la plus appropriée pour notre application [90].

Notre travail reprend la méthode de Craine *et al.* [32], en augmentant donc le nombre de capteurs mais aussi en modifiant le modèle de propagation. L'objectif principal est la vérification de la validité de cette méthode de localisation par l'introduction de la redondance et par un modèle *a priori* plus réaliste du milieu.

Modèle de propagation. Le modèle de propagation utilisé par Craine *et al.* est le modèle classique d'une source ponctuelle dans un milieu isotrope et non-absorbant. Cependant, les tissus de l'organisme humain constituent des milieux fortement absorbants pour les ondes sonores et l'abdomen n'est pas isotrope (qui plus est, son contenu varie dans le temps et donc ses caractéristiques acoustiques changent). Nous avons essayé d'améliorer le modèle de propagation par l'introduction de l'atténuation due au milieu absorbant, en sachant qu'une modélisation acoustique de l'anatomie abdominale dépasse largement nos objectifs (les premiers développements dans l'étude des propriétés acoustiques des tissus humains ont été publiés récemment par Royston, Mansy *et al.* [75, 94, 95]).

L'atténuation de l'amplitude A_i d'une onde plane dans un milieu absorbant suit une loi exponentielle [46, 52, 104] :

$$A_i = A_s e^{-\mu d_i}, \tag{4.14}$$

où μ désigne un paramètre d'absorption qui dépend des caractéristiques du milieu de propagation, mais aussi de la fréquence du son propagé, d_i désigne la distance entre le capteur i et la source et A_s l'amplitude de celle-ci. En considérant en plus l'atténuation de la puissance d'une source ponctuelle avec le carré de la distance, l'expression de l'intensité acoustique d'un son enregistré par le capteur i (4.11) se réécrit :

$$I_i = P_s \frac{e^{-2\mu d_i}}{4\pi d_i^2}. \tag{4.15}$$

Le paramètre d'absorption μ est donné par :

$$\mu = \alpha \cdot f^{\beta}, \tag{4.16}$$

où f est la fréquence du son propagé, α et β des constantes spécifiques au milieu de propagation [52]. En général, β varie entre 1 et 2 (2 pour les liquides, mais égal ou très proche de 1 pour les tissus humains). La deuxième constante α varie en fonction du tissu considéré, avec une valeur moyenne de $8,3 \cdot 10^{-6}$ (s/m) pour les tissus mous [46]. Quant à la fréquence du son propagé, nous l'avons estimée étant égale à la fréquence principale (voir éq. (4.7)) de l'événement ayant l'intensité la plus forte parmi les sons enregistrés simultanément par les voies.

En suivant un développement similaire au milieu non-absorbant, la puissance inconnue de la source peut être calculée à partir de chaque signal d'intensité I_i enregistré par le capteur i, accessible par (4.8). La position de la source est retrouvée en résolvant le système d'équations non-linéaires d'inconnues (x_s, y_s, z_s), obtenu pour toutes les paires (i, j) :

$$P_s = 4\pi d_i^2 e^{2\mu d_i} I_i = 4\pi d_j^2 e^{2\mu d_j} I_j, \tag{4.17}$$

en remplaçant les distances d_i, d_j par leurs expressions (4.9).

Remarque. L'influence de l'introduction du modèle absorbant sur la propagation est relativement peu importante. En effet, comme le paramètre d'absorption μ augmente avec la fréquence (les hautes fréquences sont plus atténuées que les basses), il est maximum pour les sons abdominaux les plus aigus. Pour $\alpha = 8,3 \cdot 10^{-6}$, $\beta = 1$ et une fréquence $f_{max} = 1250$ Hz., μ vaut 0,0104 (m^{-1}). Cette valeur indique que les sons abdominaux (qui, après la re-segmentation, ont les composantes fréquentielles supérieures à 1250 Hz. fortement diminuées) sont peu atténués sur des distances courtes, car le terme exponentiel de (4.15) atteint une valeur minimale égale à 0,993 (pour une distance maximale de propagation considérée de 1/3 m.). Autrement dit, l'influence de l'atténuation sur l'estimation des distances entre la source et les capteurs (les rayons des sphères) est de moins de 1%, donc le changement de modèle n'apporte pas une amélioration notable de la localisation. Cependant, comme ce modèle est plus général, nous allons l'utiliser pour les développements suivants.

Fonction de coût pour la localisation. La résolution du système non-linéaire (4.17) n'est pas envisageable analytiquement. De plus, l'augmentation du nombre de capteurs transforme ce système en un système sur-déterminé. Dans un milieu isotrope correctement modélisé (absorbant ou non) et en l'absence de bruit, la solution reste exacte, mais en présence de bruit de mesure et/ou si le modèle de propagation est erroné, la solution doit être recherchée par l'optimisation d'une fonction de coût $\mathcal{F}$. Avec une chaîne de mesure bien calibrée et après le débruitage appliqué antérieurement, les erreurs de mesure sont minimales, ce qui permet de considérer que la valeur minimum de $\mathcal{F}$ reflète principalement l'erreur de modélisation.

Nous présentons dans la suite la construction de cette fonction de coût. Dans le cas des erreurs de modélisation, l'équation (4.15) se réécrit :

$$\hat{P}_{s,i} = 4\pi I_i d_i^2 e^{2\mu d_i}, \tag{4.18}$$

où $\hat{P}_{s,i}$ désigne l'estimation de la puissance de la source à partir du signal enregistré par le capteur i.

Ces estimations risquent d'être différentes si le modèle est erroné. Pour vérifier leur similitude, nous pouvons soustraire deux à deux les équations du système (4.17) pour créer la fonction de coût à minimiser :

$$\mathcal{F}(x,y,z) = \sum_{\substack{i,j=1 \\ i \neq j}}^{N_v} (\hat{P}_{s,i} - \hat{P}_{s,j})^2, \tag{4.19}$$

où $N_v = 6$ est le nombre de voies (capteurs). La fonction $\mathcal{F}(x,y,z)$ est une hyper-surface positive, qui atteint son minimum global théorique en $\mathcal{F}(x_s,y_s,z_s) = 0$ quand $\hat{P}_{s,i} = \hat{P}_{s,j}$ pour toutes les permutations (i,j), ce qui est possible uniquement si le modèle de propagation (et par conséquent de localisation) est correct.

Cette fonction nécessite quelques précisions. Comme nous l'avons vu précédemment, une grande partie des sons abdominaux n'est pas détectée par tous les capteurs. L'algorithme de validation croisée permet d'éliminer les événements présents sur une seule voie ; ils sont considérés comme des artéfacts. Cependant, les sons enregistrés par un nombre de voies inférieur à 5 sont nombreux, et par conséquent les systèmes d'équations (4.13) et (4.17) sont sous-déterminés. En fait, pour les deux modèles de propagation discutés ici (4.11) et (4.15), il n'est pas possible, théoriquement, qu'un son soit détecté seulement par une partie des capteurs, car l'intensité vaut 0 uniquement à une distance infinie de la source. En pratique, bien sûr, cette hypothèse n'est pas vérifiée et, de plus, les capteurs ont une sensibilité qui leur interdit de réagir à des excitations trop faibles. Pour pallier cette difficulté, nous avons considéré que l'absence d'un son sur une ou plusieurs voies est due au seuil de sensibilité des capteurs respectifs et nous avons créé virtuellement des événements d'intensité ε_{Int} aux instants correspondants. Cette approche a l'avantage d'uniformiser l'expression de la fonction coût (4.19) pour tous les événements et de forcer une solution unique au problème de localisation.

En réalité, le minimum théorique de la fonction coût ($\mathcal{F}(x,y,z) = 0$) n'est jamais atteint. Cependant, pour chaque événement, le point de coordonnées (x_s,y_s,z_s) qui minimise cette fonction est la meilleure solution possible pour la localisation de la source du son abdominal considéré. En remplaçant ce point dans les expressions de calcul des distances d_i (4.9) et en introduisant celles-ci dans (4.18), nous obtenons les différentes estimations de la puissance de la source sonore. La dispersion de ces estimations peut être considérée comme une mesure de la validation du modèle de propagation et par conséquent de la méthode de localisation : des valeurs identiques ou proches des $\hat{P}_{s,i}$, $\hat{P}_{s,j}$ confirment la validité du modèle, alors que de grands écarts entre ces

valeurs remettent en cause les hypothèses de propagation et indiquent que la solution trouvée au problème de localisation est fausse.

Pour donner une mesure quantitative de l'adéquation du modèle, nous introduisons d'abord l'écart relatif entre l'estimation courante $\hat{P}_{s,i}$ et la moyenne de toutes les estimations (la différence normalisée entre l'estimation courante et la moyenne des estimations) :

$$\varepsilon_i = \frac{\hat{P}_{s,i} - \frac{1}{N_v}\sum_{j=1}^{N_v}\hat{P}_{s,j}}{\frac{1}{N_v}\sum_{j=1}^{N_v}\hat{P}_{s,j}}. \tag{4.20}$$

La dispersion de ces écarts, qui est une bonne mesure globale de l'adéquation du modèle, est donnée par l'erreur de modélisation ϵ :

$$\epsilon^2 = \frac{1}{N_v}\sum_{i=1}^{N_v}\varepsilon_i^2 = \frac{1}{N_v}\sum_{i=1}^{N_v}\left(\frac{\hat{P}_{s,i} - \frac{1}{N_v}\sum_{j=1}^{N_v}\hat{P}_{s,j}}{\frac{1}{N_v}\sum_{j=1}^{N_v}\hat{P}_{s,j}}\right)^2. \tag{4.21}$$

Une valeur proche de 0 de cette erreur signifie que les estimations $\hat{P}_{s,i}$ de la puissance de la source, calculées à partir des capteurs i, sont proches entre elles (et de leur moyenne). Par exemple, une erreur $\epsilon = 1$ indique un écart moyen de 100% entre les estimations $\hat{P}_{s,i}$ et leur moyenne et donc une très mauvaise adéquation du modèle de propagation aux mesures.

Résultats. Évidemment, la localisation d'un point de coordonnées (x_s, y_s, z_s) nécessite d'abord la définition d'un repère et le positionnement de la zone d'intérêt (les limites du domaine de variation de ce point) dans ce repère. Dans notre cas, l'espace des solutions possibles est l'intérieur de l'abdomen. Il est délimité en surface par les six capteurs placés sur la paroi abdominale et il a une profondeur non nulle liée à l'anatomie. L'origine du repère a été fixée comme sur la figure 4.9.

Nous avons discrétisé cet espace en unités volumiques de forme cubique, dont les dimensions, arbitraires, fixent la résolution. Ainsi, nous avons limité le domaine des solutions possibles (x_s, y_s, z_s) à un parallélépipède droit de dimensions $(7 \times 7 \times 4)$ (figure 4.9), dans lequel nous avons recherché d'une manières exhaustive le minimum de la fonction de coût $\mathcal{F}$.

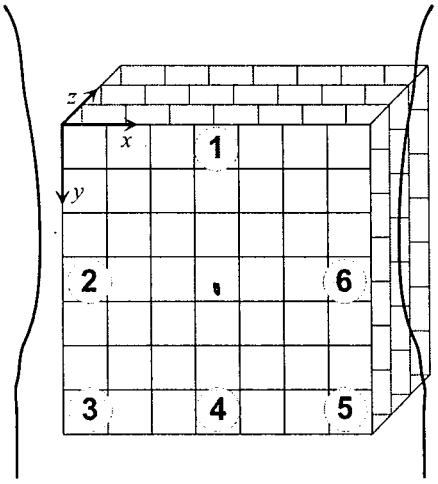


FIG. 4.9 – Repère de localisation et discrétisation de l'espace des solutions possibles (intérieur de l'abdomen).

Nous avons testé la méthode de localisation décrite précédemment sur un enregistrement de 2^{21} points, donc d'une durée d'environ 7 minutes à une fréquence d'échantillonnage de 5000 Hz.

Après les étapes d'extraction des caractéristiques physiques et de re-segmentation, nous avons recherché les événements se superposant dans le temps et nous avons détecté un total de 459 sons abdominaux distincts. Pour chacun de ces sons, nous avons calculé l'intensité acoustique (équation 4.8) et la fréquence principale (équation 4.7). Ensuite, nous avons estimé la position de la source dans le repère défini précédemment, en recherchant le minimum de la fonction coût (4.19). Pour chaque source, nous avons estimé l'erreur de modélisation ϵ (4.21). Les résultats obtenus sont décevants : l'erreur moyenne est de 1,28, avec des valeurs variant entre 0,22 et 2,45.

Pour analyser plus en détail l'erreur de modélisation, nous avons groupé les événements en fonction du nombre de capteurs les ayant détectés. L'évolution de cette erreur est présentée figure 4.10.

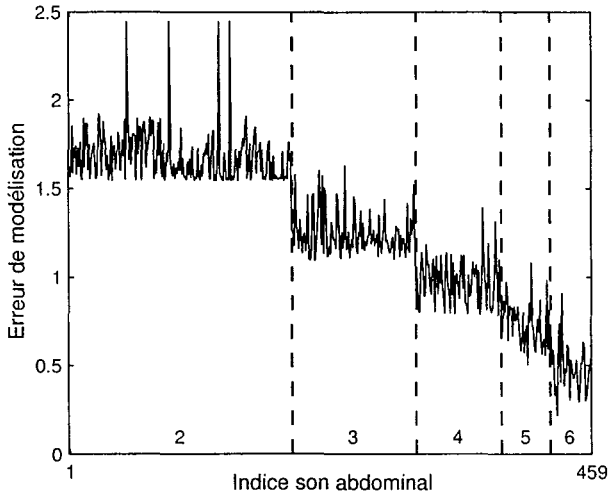


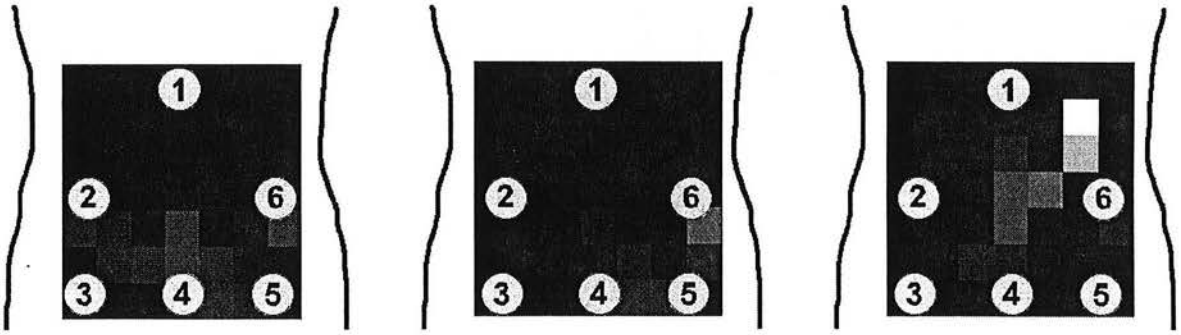
FIG. 4.10 – Exemple d'évolution de l'erreur de modélisation ϵ pour 459 sons abdominaux, groupés en fonction du nombre de voies les ayant enregistrés.

Comme on pouvait s'y attendre, l'erreur de modélisation diminue avec le nombre de capteurs : les sons les plus forts sont mieux localisés que les faibles. En fait, dans le modèle de propagation choisi, un capteur qui ne détecte pas un son le "repousse" vers l'infini. Néanmoins, même pour les sons détectés par les six capteurs (37 sur 459), l'erreur ϵ reste importante : sa valeur moyenne est de 0,49, avec un minimum à 0,22 et un maximum à 0,91. Autrement dit, même dans ce dernier cas, l'écart entre les estimées $\hat{P}_{s,i}$, $\hat{P}_{s,j}$ est proche de leur valeur moyenne.

Une autre remarque intéressante est suggérée par les valeurs des triplets (x_s, y_s, z_s) pour les événements analysés : seulement 10 sons sur 459 ont une source située dans la profondeur de l'abdomen, tous les autres étant placés sur le premier niveau ($z = 1$) du parallélépipède des solutions possibles. Le positionnement de la source sonore dans le plan des capteurs (solution qui assure des distances de propagation minimales) indique que l'augmentation de l'erreur de modélisation (et de localisation, par conséquent) est liée à l'augmentation des distances : les estimées de la puissance sonore d'une source située dans la profondeur de l'abdomen sont tellement mauvaises que la véritable localisation ne sera jamais, ou très rarement, trouvée.

Localisation empirique. Au vu des résultats précédents, on peut affirmer que la méthode de localisation des sources sonores à l'intérieur de l'abdomen en considérant un milieu de propagation isotrope (méthode mise en œuvre par Craine *et al.* [32]) est inapplicable, car les hypothèses sont loin des conditions réelles. Néanmoins, les auteurs de [32] utilisent cette méthode et les

localisations trouvées (même fausses) permettent de faire la distinction entre plusieurs maladies fonctionnelles digestives. Selon notre expérience, l'application de cette méthode (adaptée pour les 6 capteurs) permet d'observer des variations des indices d'activité au cours du temps. Par exemple, le nombre de sons par séquence de 7 minutes et par unité de volume (voir figure 4.9 pour la définition des unités de volume) varie durant les séquences successives d'un phonoentérogramme de 3 heures, comme montré figure 4.11. On peut donc affirmer que, même si la source sonore est mal localisée, l'activité abdominale détectée varie en fonction de la localisation spatiale et cette variation est mesurable si on utilise plusieurs capteurs. Autrement dit, il y a une information à extraire à partir de ces enregistrements multi-voies, et elle est intéressante du point de vue physiologique.



(a) minutes 1 à 7 (b) minutes 78 à 84 (c) minutes 155 à 161

FIG. 4.11 – Exemple de variation spatiale de l'activité abdominale au cours d'un enregistrement de longue durée. L'indice utilisé est le nombre de sons générés dans une unité de volume, définie comme sur la figure 4.9. Les zones les plus actives sont claires, alors que les zones peu actives sont sombres. La méthode de localisation (même erronée, car elle fait l'hypothèse d'un milieu de propagation isotrope) détecte des changements durant la digestion.

En auscultation clinique, les médecins déplacent le stéthoscope pour étudier plusieurs localisations. L'idée sous-jacente, implicite d'ailleurs dans la plupart des publications sur le sujet (voir par exemple [41, 48, 121]), est que les sons abdominaux s'atténuent avec la distance. Les recherches et les applications sur la propagation des sons et ultrasons dans l'organisme montrent aussi que l'abdomen est un milieu absorbant, mais fortement anisotrope [46, 52].

Dans ces conditions, il est possible de considérer un modèle de propagation moins restrictif, qui s'écrit, en utilisant les notations précédentes :

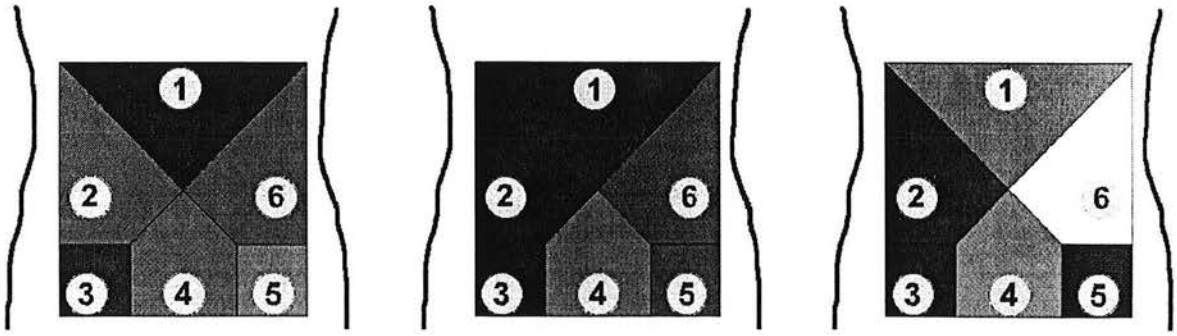
$$I_i = P_s f(d_i), \tag{4.22}$$

où f est une fonction monotone décroissante[47]. En même temps, ce modèle général est plus proche de l'expérience médicale et il est couramment et implicitement utilisé en routine clinique. Il se traduit par une règle simple, que nous avons décidé d'utiliser pour la localisation : la source d'un son abdominal se trouve dans le voisinage du stéthoscope ayant enregistré le signal le plus fort.

[47] Ce modèle, même s'il ne fait plus l'hypothèse d'un milieu isotrope, considère quand même que la fonction f (*a priori* discontinue) est la même dans toutes les directions. Cependant, il n'impose pas une loi de propagation, il est donc plus général que les précédents.

4.3.3 Résultats, exemples

La règle empirique de localisation énoncée ci-dessus implique une réduction importante de la résolution de la localisation : l'abdomen sera partagé en un nombre de zones égal au nombre de capteurs utilisés, dans notre cas six. Les résultats de cette méthode de localisation sont présentés dans la figure 4.12.



(a) minutes 1 à 7 (b) minutes 78 à 84 (c) minutes 155 à 161

FIG. 4.12 – Exemple de variation spatiale de l'activité abdominale au cours d'un enregistrement de longue durée, pour les mêmes plages que dans la figure 4.11. L'indice utilisé est le nombre des sons générés dans un volume situé dans le voisinage de chaque capteur (les zones les plus actives sont claires, alors que les zones peu actives sont sombres). La méthode de localisation utilisée est empirique.

Il est important de noter que cette méthode conduit en pratique à une nouvelle re-segmentation des signaux enregistrés. Nous avons quantifié l'importance des modifications apportées par la localisation en utilisant les indices d'activité employés pour caractériser les étapes de re-segmentation. Comme tous les événements restants après la re-segmentation sont propagés, ils sont tous affectés par la localisation. Les changements dans les indices d'activité des voies sont donc importants, comme on peut le voir dans le tableau 4.6.

Indice	Voie						Moyenne
	1	2	3	4	5	6	
N_m	2,9	5,3	3,7	18,0	11,1	14,2	9,2
D_m	0,9%	1,3%	0,6%	6,0%	3,0%	3,1%	2,5%
E_m	7881	527	351	2347	2967	6835	3485

TAB. 4.6 – Exemple de valeurs des principaux indices d'activité après localisation empirique (valeurs moyennes sur 168 minutes d'enregistrement).

A la différence des évolutions durant les étapes de re-segmentation (voir les résultats présentés dans la section 4.2.2), les variations des indices d'activité semblent liées aux six voies et donc à la localisation. Par exemple, la durée totale des événements par minute D_m est diminuée d'environ 50% après la première étape de re-segmentation pour les six voies et elle a également une évolution semblable (pour les six voies) durant les étapes suivantes (elle diminue à chaque étape entre 10 et au plus 30%, avec des différences entre les voies inférieures à 20%). En revanche, la localisation réduit l'indice D_m de presque 90% sur les voies 1 à 3, et de seulement 40% sur les voies 4 et 5, par exemple (environ 50% pour la voie 6).

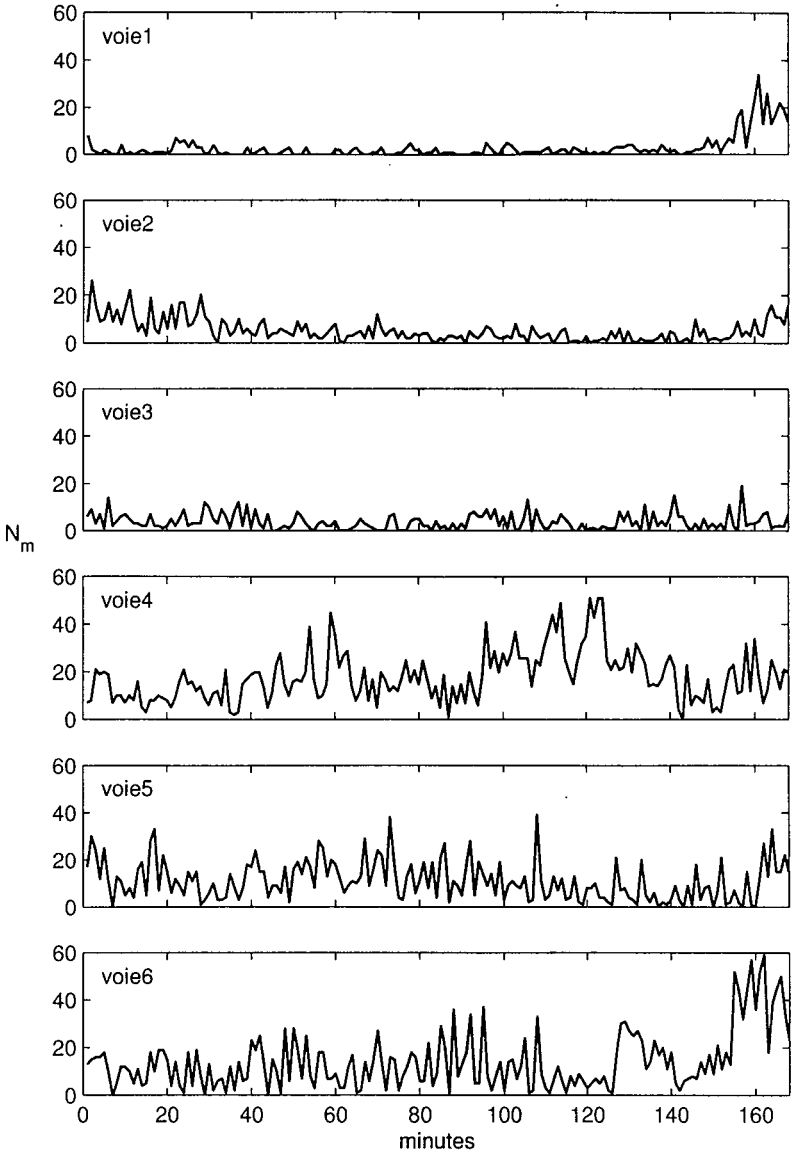


FIG. 4.13 – Évolution du nombre d'événements détectés par minute (indice N_m) par voie, durant 168 minutes.

Enfin, la localisation permet d'étudier plus en détail le fonctionnement local de l'abdomen et l'évolution au cours du temps (voir figure 4.13). Ainsi, le nombre d'événements par minute augmente pour les voies 1 et 6 vers la fin des 168 minutes d'enregistrement (les dix dernières minutes), alors qu'elle semble diminuer pour la voie no. 2 et présente des pics aux environs des minutes 60 et entre les minutes 90 et 120 pour la voie no. 4. Certes, toute interprétation est prématurée avant l'étude de plusieurs phonoentérogrammes normaux et pathologiques. De plus, il est certainement nécessaire de prendre en compte les circonstances de l'enregistrement (pour cet exemple, le pic d'activité détecté durant les dix dernières minutes est lié aux habitudes alimentaires du patient – il s'agissait de l'heure habituelle du repas).

Bien sûr, cet exemple ne constitue pas une règle, mais il montre que la localisation apporte une information supplémentaire significative dans le processus de traitement des phonoentérogrammes.

4.4 Conclusion

L'objectif déclaré de ce chapitre a été l'introduction de connaissances *a priori* dans le traitement des phonoentérogrammes, afin d'améliorer les résultats du pré-traitement précédent et de préparer l'étape d'analyse.

Nous avons d'abord présenté une méthode pour obtenir des représentations en ondelettes (des vecteurs de coefficients) pour chaque événement segmenté. Le but immédiat a été l'extraction des caractéristiques physiques intéressantes du point de vue médical telles que la durée, l'intensité acoustique et la fréquence principale. L'intégration de ces connaissances dans la méthodologie de traitement des phonoentérogrammes est faite à l'aide d'un algorithme de re-segmentation et d'élimination d'artéfacts (section 4.2.1). Certes, ces trois caractéristiques physiques peuvent être obtenues après une reconstruction directe du signal. Cependant, comme l'objectif principal de cette thèse est la création d'un ensemble de méthodes et d'algorithmes pour le traitement de sons physiologiques, nous considérons que la description précise de chaque événement, offerte par sa représentation en ondelettes, mérite d'être extraite au vue d'une analyse ultérieure.

Nous avons également introduit et défini plusieurs indices d'activité (eux aussi recensés dans la littérature), tels que le nombre d'événements par minute, la durée totale des événements par minute etc., qui permettent d'apprécier quantitativement la motricité des organes de l'abdomen et d'évaluer l'effet de l'introduction de connaissances *a priori*.

Il faut souligner aussi l'importance de l'approche multi-voies. L'information apportée est introduite, grâce à la validation croisée, comme une dernière étape de l'algorithme de re-segmentation. Cependant, l'apport principal des enregistrements multi-voies est la possibilité qu'ils offrent de localiser les sources des sons abdominaux. Nous avons abordé le problème de la localisation à partir des références bibliographiques (peu nombreuses) dans le domaine, en cherchant une validation systématique nécessaire à toute application. Les tests montrent que les méthodes de localisation trop simplifiées, notamment celles qui considèrent l'abdomen comme un milieu de propagation isotrope, sont inadaptées. Notre opinion est que les hypothèses faites sont erronées et que l'introduction d'autres méthodes de localisation (comme celles prenant en compte l'intercorrélation des voies) ou d'autres fonctions de coût (non quadratiques, par exemple) ne changera pas cette conclusion. Nous proposons donc un modèle empirique, mais néanmoins justifié, et nous présentons les résultats obtenus sur un exemple de phonoentérogramme réel.

Les traitements présentés dans ce chapitre permettent de conclure que :

- les connaissances *a priori* sur la nature des signaux sont essentielles pour avancer vers une application pratique. Ces connaissances, liées aux caractéristiques physiques des sons abdominaux telles que la durée, la fréquence principale ou l'intensité acoustique, peuvent être extraites directement de la décomposition en ondelettes de chaque événement, accessible après l'étape de débruitage ;
- l'approche multi-voies apporte une information importante car, même si on ne dispose pas d'une méthode de localisation précise, on peut détecter les variations locales de l'activité abdominale. L'exemple présenté ne permet pas de valider cette approche, mais suggère néanmoins l'utilité d'une investigation et d'une validation plus approfondies.



Analyse des phonoentérogrammes

Les travaux présentés dans les chapitres précédents (le débruitage, la segmentation, la re-segmentation par l'élimination des artéfacts et la localisation) peuvent être regroupés sous la dénomination générique de **pré-traitement**. Afin de les mettre en œuvre et d'évaluer leurs effets sur les signaux sonores abdominaux, nous avons été amenés à extraire des caractéristiques physiques (4.1.1) qui décrivent les sons abdominaux individuellement. Nous avons défini plusieurs indices d'activité (4.1.2) qui caractérisent globalement les phonoentérogrammes. A l'aide de ces indices et des caractéristiques physiques, nous abordons dans ce chapitre **l'analyse multidimensionnelle** des signaux et des événements, qui servira ensuite à l'interprétation et au diagnostic médical (voir le schéma d'aide à la décision présenté dans l'introduction, figure 1.2, page 17).

Les résultats que nous présentons dans ce chapitre ont essentiellement une fonction d'exemple. Nous n'avons ni l'ambition ni la capacité de donner une véritable interprétation physiologique de ces résultats, et l'objectif principal est de présenter des méthodes d'analyse et non pas les résultats de cette analyse.

Plusieurs publications dans le domaine de la phonoentérographie proposent des méthodes d'analyse globale des signaux abdominaux. Cependant, à notre connaissance, aucune n'aborde le problème dans un espace multidimensionnel : dans la littérature, on étudie les signaux en fonction de chaque indice d'activité pris séparément, mais jamais en fonction de plusieurs indices simultanément (voir section 1.1.2). Des interprétations physiologiques ou médicales simples et parfois contestables (voir les travaux de Craine sur les maladies fonctionnelles [30, 31], contestés par Yuki [127]) sont alors avancées.

Une autre approche, plus difficile et rarement abordée dans la bibliographie, est l'analyse et la classification des événements extraits du signal mesuré. L'intérêt de cette approche est double :

- elle permet d'identifier des sons spécifiques à certaines pathologies (l'occlusion intestinale par exemple [49]) ;
- elle permet d'étudier plus en détail le fonctionnement des organes abdominaux, en associant par exemple un certain type de son à un organe particulier et/ou à un stade de la digestion. Comme pour les valeurs des indices d'activité, cette mise en correspondance type de son–organe générateur–stade de digestion peut être ensuite étudiée sur des sujets sains, mais

également dans des conditions physiologiques liées aux différentes maladies ou traitements (médicamenteux ou chirurgicaux).

Malheureusement, la plupart des descriptions des sons abdominaux dans la littérature sont subjectives et couvrent des classes vaguement définies : les gargouillements, ou borborygmes, sont caractérisés éventuellement par leurs caractéristiques physiques (durée, fréquence), mais plus souvent par des "métaphores" (sons explosifs, sons métalliques, *clicks*, *staccato pops* etc.). Comme dans le cas des indices d'activité, aucune analyse prenant en compte plusieurs caractéristiques simultanément n'a été proposée.

L'objectif de ce chapitre est d'aborder l'analyse (globale et locale) des phonoentérogrammes dans un espace multidimensionnel. Dans ce cadre, les données (les sons individuels et les séquences de phonoentérogramme) sont considérées comme des entités décrites par leurs variables explicatives respectives (caractéristiques physiques et indices d'activité). La représentation d'une entité est un point dans l'espace engendré par les variables explicatives qui constituent ses axes. Les caractéristiques de cet espace dépendent des variables explicatives utilisées, de leur nombre (qui détermine la dimension de l'espace), de leurs valeurs (qui servent à calculer la distance entre deux points, selon la métrique utilisée) et des relations entre elles (variables corrélées ou non corrélées).

Dans cet espace, on cherche à trouver des ensembles homogènes de signaux/événements (des "nuages de points"). La difficulté principale de cette démarche, dans le domaine de la phonoentérographie, est l'absence de toute classification prédéfinie : nous ne disposons pas d'une collection de sons ou de signaux type, spécifiques à certains phénomènes physiologiques ou physiopathologiques.

Ce chapitre est structuré en deux grandes parties :

- la première est dédiée à l'analyse en composantes principales, avec ses deux volets :
  - l'étude des caractéristiques et des espaces engendrés pour la représentation des événements individuels et des séquences de phonoentérogramme. Par caractéristiques nous désignons à la fois les caractéristiques physiques, dans un contexte concernant les sons individuels, et les indices d'activité, quand l'objectif est l'analyse globale des phonoentérogrammes. Cette étape aboutit à la création de deux espaces de représentation (pour les séquences de phonoentérogramme, respectivement pour les événements) ;
  - l'interprétation des résultats de l'ACP dans les espaces définis précédemment. Nous appliquons cette méthode sur plusieurs phonoentérogrammes réels et nous présentons les résultats ;
- la seconde partie étudie les possibilités d'analyse offertes par les méthodes de reconnaissance des formes, en particulier par la classification non-supervisée.

5.1 Analyse en composantes principales

5.1.1 Espaces de caractéristiques

La première étape de l'analyse (globale ou locale) des phonoentérogrammes commence par le choix des caractéristiques adaptées et la construction de l'espace de représentation. Il est important de noter que la pertinence d'un jeu de caractéristiques utilisé ensuite pour l'analyse des données (la classification) ne peut pas être appréciée avant sa validation par les résultats, les frontières qui séparent ces deux étapes étant plutôt arbitraires [40]. Autrement dit, tous les couples constitués par un jeu de caractéristiques et un algorithme de classification devraient être testés et comparés avant un choix définitif. Cette tâche dépasse les objectifs immédiats de notre travail : notre but principal est de vérifier si l'étude des sons abdominaux est informative pour

l'évaluation de l'activité gastro-intestinale et de fournir des outils pour l'appréciation quantitative et objective de cette activité. Ces outils ne seront sans doute pas les plus performants, mais une fois validés, leur amélioration ultérieure est envisageable et souhaitable.

Selon Celeux [21], on distingue trois types de caractéristiques (variables explicatives) :

- les caractéristiques utiles, qui sont spécifiques à certaines classes ;
- les caractéristiques redondantes, dont l'information discriminante est contenue dans d'autres variables ;
- les caractéristiques nuisibles, qui augmentent inutilement la dimension de l'espace de classification et qui sont distribuées uniformément dans toutes les classes.

Une bonne description d'un signal découle de l'emploi de variables explicatives utiles. Elle doit être envisagée sous deux aspects [40] : d'une part, elle doit offrir la possibilité de distinguer des signaux (événements) différents, et d'autre part, elle doit avoir la capacité de reconnaître des signaux (événements) semblables, autrement dit de les classifier. De plus, si on considère l'étape d'interprétation qui suit la classification, les caractéristiques choisies doivent fournir une image pertinente, complète et correcte du signal analysé.

La construction d'un espace de représentation pertinent à partir du jeu de caractéristiques revient à rechercher les caractéristiques utiles, ou encore à éliminer de la redondance et les variables nuisibles. Le problème peut être abordé sous deux angles [36] :

- **l'extraction des caractéristiques** crée de nouvelles variables explicatives non-corrélées et cherche à garder un maximum de variance possible dans l'ensemble des points à classifier. Une méthode classique est l'analyse en composantes principales (ACP) (voir par exemple [40, 100]), qui redéfinit les axes de l'espace de classification (les nouvelles variables explicatives) comme des combinaisons linéaires des anciennes caractéristiques[48], tout en leur assurant l'orthogonalité et des directions privilégiées qui maximisent la variance. Un avantage important de l'ACP est la possibilité de réduction du nombre de variables explicatives et donc de la dimension de l'espace de travail : les nouvelles variables (composantes principales) sont ordonnées selon leur variance, et l'utilisateur peut choisir les P premières composantes et créer ainsi un sous-espace de classification qui garde l'essentiel de la variance du nuage de points. L'inconvénient principal de la méthode est le manque de signification physique des nouvelles variables explicatives. Cependant, des techniques comme le cercle de corrélation [100] peuvent faciliter la compréhension et l'interprétation des résultats.
  Il n'y a pas de raison d'affirmer qu'un axe maximisant la variance est le plus adapté pour faciliter l'analyse des données et leur séparation en classes. Cette séparation est le but de l'analyse discriminante [21, 40, 100], qui vise aussi à trouver de nouveaux axes de l'espace de représentation. A la différence de l'ACP, l'analyse discriminante cherche les axes qui ont un pouvoir de séparation maximum. Cependant, elle nécessite un ensemble de points déjà séparés, appelé ensemble d'apprentissage, afin de trouver les axes respectifs, ce dont nous ne disposons pas. L'analyse discriminante est surtout appliquée dans les problèmes de classification supervisée, comme nous allons le montrer à la fin de ce chapitre ;
- **la sélection des caractéristiques**, comme son nom l'indique, implique le choix des variables explicatives utiles, pertinentes. Selon Blum et Langley [13], la question importante est : "pertinentes pour quoi" ? Effectivement, la sélection des variables explicatives utiles parmi les caractéristiques des signaux à classifier est souvent abordée, comme l'analyse discriminante, dans le cadre de la classification supervisée, c'est-à-dire quand on dispose d'un ensemble d'apprentissage déjà classé. Grâce à cet étalon, on peut vérifier et com-

[48] Afin de leur donner une importance égale et de prendre en compte les différences entre les valeurs numériques dues aux unités de mesure, les caractéristiques sont normalisées avant l'ACP.

parer les sous-ensembles des caractéristiques sélectionnées (voir par exemple [13, 21, 35]) afin de choisir les variables utiles. La comparaison peut se faire de plusieurs manières, plus ou moins optimales, mais disposant toutes d'un critère de validation direct : le jeu de caractéristiques choisi sépare-t-il correctement l'ensemble d'apprentissage en les classes prédéfinies ? Malheureusement, comme nous l'avons précisé précédemment, ce n'est pas notre cas : nous ne disposons d'aucune classification objective *a priori* des sons ou des signaux abdominaux.

Le choix initial des variables explicatives porte sur des caractéristiques empiriques, c'est-à-dire les caractéristiques employées pour la description médicale des signaux d'intérêt : il s'agit des indices d'activité (voir section 4.1.2), utilisés pour la description globale des séquences de phonoentérogramme, et des caractéristiques physiques (voir section 4.1.1), pour la description locale des sons abdominaux pris individuellement. Cette solution, même si elle n'est pas rigoureuse (car on pourrait envisager un choix initial des variables explicatives basé sur la représentation en ondelettes des signaux), a l'avantage d'être validée par l'expérience (médicale, dans notre cas) et assure une relative pertinence des variables explicatives sélectionnées.

Cependant, les différentes sources bibliographiques ne s'accordent pas sur les caractéristiques à utiliser pour la description des signaux. Cette observation est surtout valable dans le cas des indices d'activité utilisés pour la description globale des phonoentérogrammes. Par conséquent, nous considérons nécessaire une étude de l'ensemble des indices présentés dans la littérature afin d'éliminer les redondances et de construire un espace de représentation adéquat. A notre connaissance, seuls Craine *et al.* [31] étudient la redondance des différents indices, pour trouver de fortes corrélations entre le nombre de sons par minute (N_m) et la durée totale des sons par minute (D_m), anti-corrélés avec la durée moyenne de l'intervalle de silence entre deux sons ($D_{s,\mu}$).

Concernant l'analyse des sons abdominaux individuels, la description médicale la plus précise est axée sur les trois caractéristiques physiques principales : la durée, la fréquence principale et l'intensité acoustique. Ces caractéristiques physiques sont d'ailleurs souvent utilisées pour l'analyse des signaux sonores en général, car elles sont proches de ce qu'on appelle caractéristiques perceptuelles (*perceptual features*), comme le *pitch* ou la *loudness* [50]. D'autres variables explicatives, toujours avec une signification physique, peuvent être employées : il s'agit de la puissance d'un son ou de son énergie, ou encore du nombre de passages par zéro (lié à la fréquence, dans le cas des sons à contenu fréquentiel pauvre).

Nous proposons dans la suite de cette section d'étudier les redondances possibles entre les variables explicatives sur un exemple de phonoentérogramme réel, afin de construire les espaces de classification adéquats. Nous abordons successivement le cas des séquences de phonoentérogramme, décrites par les indices d'activité, et celui des sons abdominaux individuels (événements), décrits par leurs caractéristiques physiques.

Indices d'activité

Les indices d'activité considérés sont ceux définis dans la section 4.1.2 (l'unité de temps est la minute) :

1. le nombre d'événements N_m ;
2. l'énergie totale E_m ;
3. la durée totale des événements (en pour-cent) D_m ;
4. l'énergie moyenne des événements E_μ ;
5. la durée moyenne des événements D_μ ;
6. la puissance moyenne des événements P_μ ;

7. la fréquence principale moyenne des événements f_μ ;
8. l'intensité acoustique moyenne des événements I_μ ;
9. la durée moyenne des silences séparant les événements $D_{s,\mu}$.

A titre d'exemple, l'évolution au cours du temps des neuf indices d'activité calculés sur une voie (no. 4) est présentée figure 5.1. Un autre exemple, concernant les valeurs prises sur les 6 voies par un seul indice (le nombre d'événements par minute N_m) a été présenté figure 4.6, page 97.

Nous considérons comme signal de travail le même phonoentérogramme que nous avons utilisé pour tester les algorithmes de re-segmentation et de localisation du chapitre précédent. Il s'agit d'un enregistrement sur 6 voies d'environ 168 minutes (12 séquences à 2^{22} points, à une fréquence d'échantillonnage de 5000 Hz.). Les données (les individus) sont des séquences d'une minute, pour lesquelles nous avons calculé les neuf indices d'activité énumérés ci-dessus. Nous obtenons donc un ensemble de $168 \cdot 6 = 1008$ points dans un espace 9-dimensionnel. L'ensemble des données constitue une matrice $\mathbf{D}$ de dimension 1008×9.

L'interprétation directe des 9 courbes présentées figure 5.1 est difficile. Toutefois, on aperçoit certaines corrélations entre les indices. Par exemple, l'évolution du nombre d'événements par minute N_m semble liée à celle de la durée totale des sons D_m. L'énergie totale par minute E_m semble aussi corrélée à l'énergie moyenne des événements E_μ. Une remarque plus intéressante peut être faite à propos de la durée moyenne des événements D_μ et la durée totale D_m : ces indices semblent assez corrélés dans la première partie de l'enregistrement, mais la durée totale augmente entre les minutes 90 et 130 sans être suivie par l'augmentation de la durée moyenne des événements (au contraire, elle est corrélée avec le nombre d'événements). Une interprétation possible est le changement du type d'activité entre les minutes 70 et 90 : si l'on avait, entre les minutes 30 et 70, une activité qui se manifestait par des événements de plus longue durée, on a des sons plus nombreux mais plus courts entre les minutes 90 et 130.

Un problème important, outre les difficultés d'interprétation, est que les variations des indices d'activité présentées figure 5.1 sont plutôt spécifiques à ce phonoentérogramme, les différences inter-patients étant importantes. Cette observation suggère qu'il est peut-être utile de traiter en même temps plusieurs phonoentérogrammes, afin de moyenner les différences et d'extraire l'éventuelle information contenue. L'interprétation risque de devenir encore plus difficile quand l'expert doit prendre en compte une quantité trop grande d'information.

Une solution est l'analyse en composantes principales, que nous appliquons à l'ensemble des données disponibles. Pour l'instant, il s'agit d'un nombre réduit d'enregistrements – 3 phonoentérogrammes multi-voies de 168 minutes chacun, donc un total de 3024 séquences d'une minute, caractérisées par les 9 indices d'activité. Pour réduire la dimension de l'espace et pour dévoiler les redondances éventuelles entre les neuf variables explicatives, nous avons calculé leur matrice de corrélation (les coefficients, multipliés par 100 et arrondis, sont présentés tableau 5.1).

Une première lecture de la matrice de corrélation (tableau 5.1) confirme le constat antérieur sur le lien entre le nombre d'événements par minute N_m et la durée totale des événements par minute D_m, ainsi qu'une corrélation relativement importante entre l'énergie par minute E_m et l'énergie moyenne des événements par minute E_μ. Des corrélations moins importantes sont visibles entre l'intensité acoustique moyenne des événements I_μ et leur puissance moyenne P_μ, ainsi qu'entre ce dernier indice et l'énergie moyenne E_μ. Au contraire, les durées moyennes des événements D_μ et des silences $D_{s,\mu}$ semblent peu corrélées avec les autres variables explicatives, donc leur potentiel informatif mérite d'être étudié. Tous ces constats confirment l'utilité de la réduction de la dimension de l'espace de données par l'élimination de la redondance.

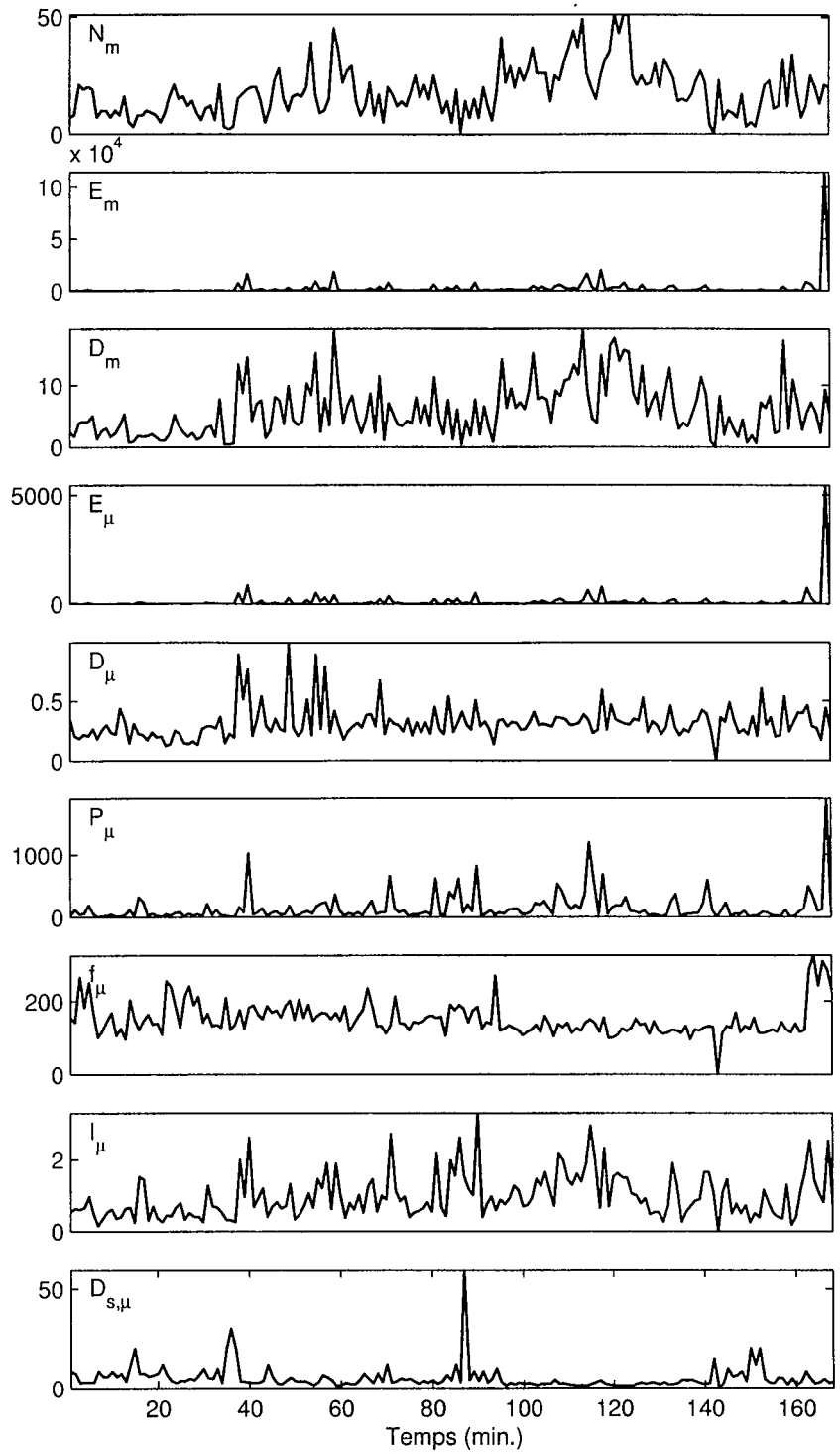


FIG. 5.1 – Exemple de variations des indices d'activité pour la voie numéro 4. Les valeurs numériques des indices varient d'une manière importante d'un indice à l'autre.

La méthode la plus utilisée est l'analyse en composantes principales. La présentation et l'interprétation des résultats de l'ACP suit la démarche indiquée par Saporta [100].

	N_m	E_m	D_m	E_μ	D_μ	P_μ	f_μ	I_μ	$D_{s,\mu}$
N_m	100	20	81	5	9	15	17	17	-48
E_m	20	100	40	72	24	64	7	29	-8
D_m	81	40	100	30	49	29	6	37	-41
E_μ	5	72	30	100	52	61	4	38	-4
D_μ	9	24	49	52	100	26	5	50	-3
P_μ	15	64	29	61	26	100	19	72	-10
f_μ	17	7	6	4	5	19	100	31	11
I_μ	17	29	37	38	50	72	31	100	-13
$D_{s,\mu}$	-48	-8	-41	-4	-3	-10	11	-13	100

TAB. 5.1 – Matrice de corrélation des indices d'activité.

L'ACP crée un système d'axes orthogonaux, appelés axes principaux, et re-projette les données sur ces axes. On obtient ainsi une nouvelle matrice de données **M** ayant la même dimension que la matrice d'origine **D**. Chaque colonne de **M** représente une nouvelle variable explicative, appelée "composante principale". A la différence des variables explicatives d'origine (les colonnes de **D**), les composantes principales (les colonnes de **M**) sont orthogonales et ordonnées selon leurs variances, en ordre décroissant. Ces variances permettent d'estimer l'importance des nouvelles variables explicatives, et de choisir par la suite les axes principaux qu'on souhaite garder dans le nouvel espace de caractéristiques réduit. Le tableau 5.2 présente les variances de chaque nouvelle variable, ainsi que le pourcentage de la variance totale expliquée et le cumul de ces pourcentages. Ce tableau aide à choisir les axes (les variables) significatifs : comme la réduction de l'espace revient à re-projeter les données sur un sous-espace de dimension réduite, il est souhaitable que ce sous-espace garde l'essentiel de l'information, qu'il ne déforme pas le nuage de points initial. Autrement dit, on doit garder un maximum de variance après re-projection.

	var	%	% cumulé
c_1	3,51	39,0	39,0
c_2	1,76	19,6	58,6
c_3	1,16	12,9	71,5
c_4	0,94	10,4	81,9
c_5	0,76	8,5	90,3
c_6	0,50	5,4	95,7
c_7	0,20	2,3	98,0
c_8	0,11	1,3	99,3
c_9	0,06	0,7	100,0

TAB. 5.2 – Variances des composantes principales de l'espace des indices d'activité et leurs contributions, en pour-cent, à la variance totale des données.

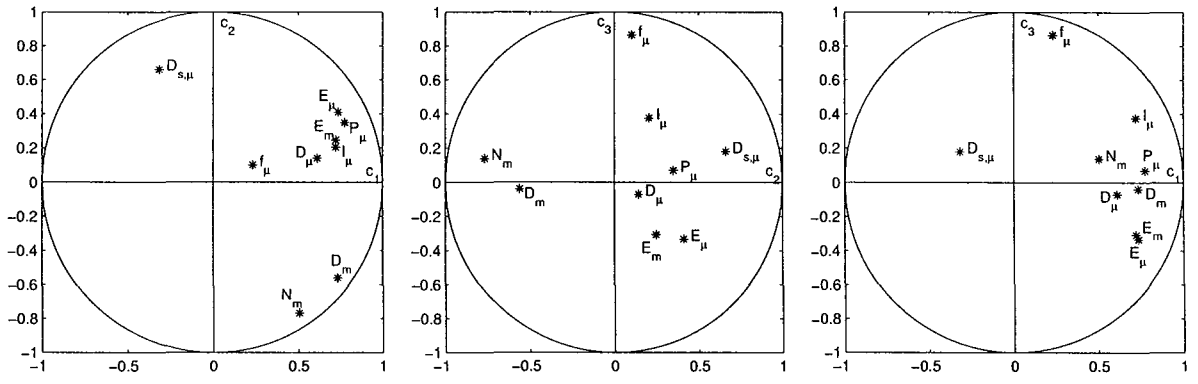
C'est à l'utilisateur de choisir le critère de sélection des axes principaux. Selon Saporta [100], le plus connu est le critère de Kaiser, qui préconise le maintien des composantes de variance supérieure à 1, ce qui, dans notre cas, revient à garder les trois premiers axes. L'espace tridimensionnel ainsi créé garde plus de 70% la variance initiale (71,5 %, voir tableau 5.2). Une autre

méthode, elle aussi empirique, propose comme critère de limitation de la dimension de l'espace un seuil de 80% de la variance, ce qui reviendrait à garder les quatre premières composantes principales. Néanmoins, il reste l'obligation de ne retenir que les composantes interprétables, c'est-à-dire les composantes auxquelles on peut donner une signification physique grâce à leurs corrélations avec les variables explicatives d'origine [100].

Les indices de corrélation entre les nouvelles colonnes (de la matrice **M**) et les colonnes de la matrice **D** sont les corrélations entre les nouvelles variables explicatives obtenues par ACP et les caractéristiques d'origine. Les relations entre ces deux variables (tableau 5.3) sont habituellement représentées à l'aide des cercles de corrélation (figure 5.2). La construction de ces cercles est intuitive : les axes des abscisses et des ordonnées représentent les composantes principales, et les variables explicatives d'origine sont des points ayant pour abscisse et ordonnée les corrélations avec les composantes principales respectives.

	c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8	c_9
N_m	50	-77	14	-12	22	-16	-16	-2	14
E_m	72	25	-31	-40	27	2	28	-5	6
D_m	73	-56	-4	18	22	-17	6	2	-18
E_μ	74	41	-33	-4	15	25	-28	-9	-4
D_μ	61	14	-7	74	7	15	7	11	8
P_μ	78	35	7	-31	-28	-20	-8	22	0
f_μ	23	10	86	-16	25	32	3	3	-2
I_μ	72	21	38	18	-43	-19	4	-20	1
$D_{s,\mu}$	-32	66	18	16	48	-41	-6	-2	0

TAB. 5.3 – Matrice de corrélation entre les composantes principales et les variables explicatives d'origine (les indices d'activité). Les coefficients sont multipliés par 100 et arrondis.



(a) axes principaux c_1–c_2 (b) axes principaux c_2–c_3 (c) axes principaux c_1–c_3

FIG. 5.2 – Cercles des corrélations indices d'activité – composantes principales, pour les trois premiers axes principaux.

A partir du tableau 5.3 et des cercles de corrélations (figure 5.2), nous pouvons interpréter la signification des composantes principales. Ainsi, la première colonne de la matrice des corrélations

(tableau 5.3) et les positions des points dans les cercles des corrélations (figures 5.2(a) et 5.2(c)) montrent une corrélation relativement importante entre la nouvelle variable c_1 et la puissance moyenne des événements P_μ, mais aussi avec l'énergie moyenne E_μ, l'énergie totale E_m, l'intensité acoustique moyenne des événements I_μ (ces caractéristiques sont d'ailleurs assez corrélées entre elles, voir tableau 5.1). Cette composante est aussi corrélée avec la durée totale des sons par minute D_m, ce qui est assez surprenant. En effet, comme c_1 est corrélée positivement avec presque tous les indices d'activité, on peut l'interpréter comme une variable "de taille" [100] : les points qui ont des valeurs importantes de cette variable ont en effet des valeurs relativement importantes pour tous les indices d'activité, ils ont une grande "taille". Néanmoins, la première composante principale est surtout reliée aux indices de type énergétique. Par conséquent, nous proposons son interprétation comme un "niveau sonore" des événements apparaissant durant la minute : une valeur importante de cette nouvelle variable explicative signale que la minute étudiée est active du point de vue niveau sonore. Cette composante explique 39,5% de la dispersion initiale (tableau 5.2).

La deuxième composante principale c_2 est anti-corrélée avec le nombre d'événements par minute N_m (figure 5.2(b)) et assez corrélée avec la durée moyenne des silences $D_{s,\mu}$. Cette observation suggère une interprétation de c_2 comme une mesure de la pauvreté en événements de la minute étudiée : plus la valeur de la nouvelle variable est élevée, plus la minute est pauvre en événements et silencieuse[49] (d'ailleurs, c_2 est aussi anti-corrélée avec D_m). Cette composante explique 19,6% de la dispersion du nuage des points (tableau 5.2).

La troisième composante principale, qui explique 12,9% de la variance totale, est clairement liée à la fréquence moyenne des événements, comme présenté figure 5.2(c) : une valeur élevée signifiera que les événements de cette séquence sont plutôt haute fréquence par rapport à la moyenne.

Par conséquent, un espace réduit représentant les données (les séquences à analyser) peut être construit à partir des trois premières composantes principales. Il maintient 71,5% de la variance initiale du nuage de points, et ses trois axes (orthogonaux) ont des interprétations physiques intéressantes et *a priori* utiles pour l'application : une minute (ou toute autre partie de phonoentérogramme) sera décrite par son niveau sonore (c_1), par sa pauvreté en événements (c_2) et par sa fréquence moyenne (c_3). A cet espace on peut ajouter un quatrième axe (c_4), qui apporte l'explication de 10,4% supplémentaires de la variance initiale et qui, de plus, est interprétable physiquement : la colonne correspondante du tableau 5.3 montre une corrélation assez importante avec l'indice D_μ (durée moyenne des événements). Le choix devra être fait après une étape de validation de l'analyse et d'interprétation des résultats. Néanmoins, pour la facilité de la représentation, nous allons nous limiter pour l'instant aux trois premières composantes principales. Dans l'espace engendré, le nuage de 3024 points a l'aspect présenté figure 5.3.

Caractéristiques physiques

La construction de l'espace de représentation des événements (sons abdominaux individuels) suit la même démarche que celle présentée pour l'espace adapté aux phonoentérogrammes. Les variables explicatives de départ (les indices d'activité dans le cas des phonoentérogrammes) sont les caractéristiques physiques suivantes, au nombre de cinq :

1. la durée d'un événement D_e ;
2. l'énergie de l'événement E_e ;

[49] Une observation similaire sur l'anti-corrélation de ces indices a été faite par Craine *et al.* [30, 31].

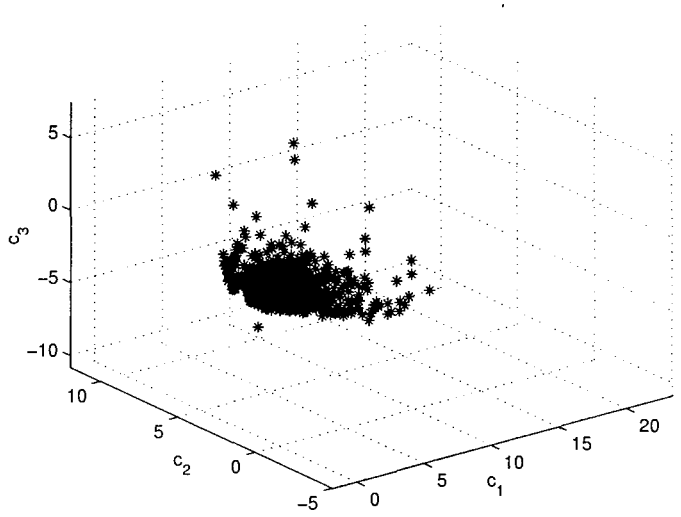


FIG. 5.3 – Répartition 3D des 3024 minutes d'enregistrement dans l'espace engendré par les 3 premières composantes principales.

3. la puissance $P_e = \dfrac{E_e}{D_e}$;
4. l'intensité acoustique I_e ;
5. la fréquence principale f_e.

Nous considérons comme signal de travail les mêmes phonoentérogrammes – trois enregistrements sur six voies d'environ 168 minutes ($24 \cdot 2^{21}$ points à une fréquence d'échantillonnage de 5000 Hz.). Après les étapes de pré-traitement, nous obtenons 32790 événements, décrits par cinq variables explicatives, ce qui revient à considérer un nuage de 32790 points dans $\mathbb{R}^5$.

Les données sont regroupées dans une matrice $\mathbf{E}$ de dimension 32790×5. Comme dans le cas des indices d'activité, nous avons calculé leur matrice de corrélation (les coefficients, multipliés par 100 et arrondis, sont présentés tableau 5.4) et nous avons effectué une analyse en composantes principales.

	D_e	E_e	P_e	I_e	f_e
D_e	100	32	18	44	-5
E_e	32	100	76	28	3
P_e	18	76	100	54	8
I_e	44	28	54	100	18
f_e	-5	3	8	18	100

TAB. 5.4 – Matrice de corrélation des caractéristiques physiques des événements.

La lecture de la matrice de corrélation (tableau 5.4) ne montre pas de corrélations très fortes entre les variables explicatives, sauf, peut-être, entre l'énergie E_e d'un événement et sa puissance P_e.

L'analyse en composantes principales crée de nouvelles variables explicatives, représentées par des axes orthogonaux. Leur importance et leur pouvoir explicatif sont donnés par les variances des projections sur ces axes, présentées dans le tableau 5.5.

	var	%	% cumulé
c_1	2,30	46,1	46,1
c_2	1,05	20,9	67,0
c_3	0,94	18,9	85,9
c_4	0,57	11,5	97,4
c_5	0,13	2,6	100,0

TAB. 5.5 – Variances des composantes principales de l'espace des caractéristiques physiques et leurs contributions, en pour-cent, à la variance totale des données.

	c_1	c_2	c_3	c_4	c_5
D_e	57	34	65	36	10
E_e	81	13	-42	33	-20
P_e	87	-5	-40	-17	24
I_e	75	-15	40	-49	-14
f_e	16	-94	13	26	2

TAB. 5.6 – Matrice de corrélation entre les composantes principales et les variables explicatives d'origine (les caractéristiques physiques). Les coefficients sont multipliés par 100 et arrondis.

Les trois premiers axes créent un espace tridimensionnel qui garde environ 86% de la variance initiale. Les significations physiques de ces axes (les trois premières composantes principales) sont obtenues à l'aide des corrélations entre les nouvelles variables explicatives obtenues par ACP et les caractéristiques d'origine. La matrice des corrélations est présentée tableau 5.6, et les cercles des corrélations figure 5.4.

A partir du tableau 5.6 et des cercles des corrélations (figure 5.4), nous pouvons interpréter la signification des composantes principales. De nouveau, la première composante a un effet global de taille : les événements qui sont caractérisés par des valeurs importantes de cette variable sont en effet des événements à la fois plus énergétiques, plus longs et de (légèrement) plus haute fréquence que la moyenne. Plus précisément, la première colonne de la matrice de corrélations 5.6 et les positions des points dans les cercles (figures 5.4(a) et 5.4(c)) montrent une corrélation importante entre la nouvelle variable c_1 et la puissance des événements P_e, l'énergie E_e et l'intensité acoustique I_e. Comme pour les indices d'activité, nous pouvons relier la première composante principale au "niveau sonore" des événements : une valeur importante de cette nouvelle variable explicative, qui a une signification énergétique, signale que l'événement étudié a un niveau sonore important. Cette composante explique 46,1% de la dispersion du nuage des points (tableau 5.5).

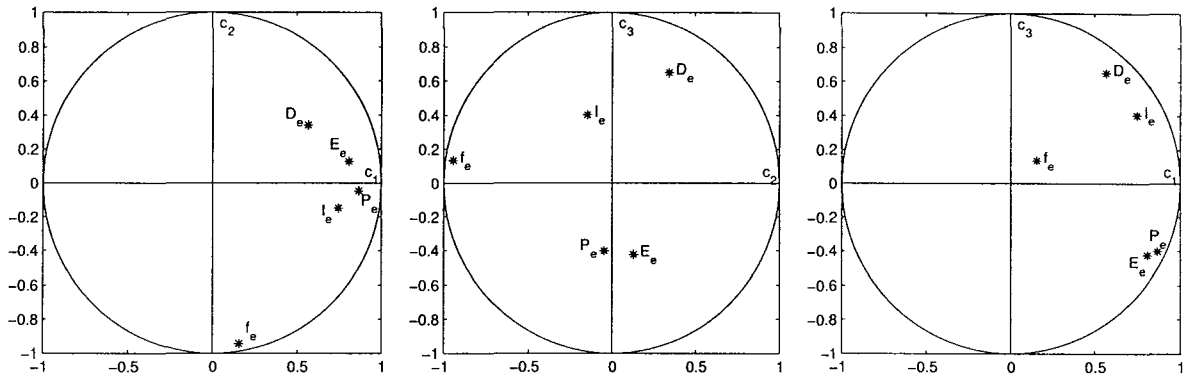
La deuxième composante principale c_2 est fortement anti-corrélée à la fréquence principale des événements f_e (figures 5.4(a) et 5.4(b)) : plus la valeur de la nouvelle variable explicative est élevée, plus la fréquence de l'événement est basse. Cette composante explique 20,9% de la dispersion du nuage de points (tableau 5.5).

La troisième composante principale, qui explique 18,9% de la variance totale, est assez corrélée à la durée des événements D_e, comme présenté figure 5.4(b) et tableau 5.6 : plus sa valeur est importante, plus l'événement est long.

Un espace réduit représentant les données (les événements) peut être construit à partir des trois premières composantes principales. Il maintient presque 86% de la variance initiale du nuage de points, et ses trois axes (orthogonaux) ont des interprétations physiques : un événement sera décrit par son niveau sonore (c_1), par sa fréquence (c_2) et, d'une manière moins évidente, par une variable reliée à sa durée (c_3).

5.1.2 Analyse des données selon l'ACP

Le rôle principal de l'ACP est la création d'un espace de représentation efficace : il préserve en bonne partie les propriétés initiales du nuage des points dans un espace de dimension réduite. Les axes de ce nouvel espace sont de nouvelles variables explicatives qui peuvent être interprétées physiquement à l'aide des cercles de corrélation.



(a) axes principaux c_1–c_2 (b) axes principaux c_2–c_3 (c) axes principaux c_1–c_3

FIG. 5.4 – Cercles des corrélations caractéristiques physiques – composantes principales, pour les trois premiers axes principaux.

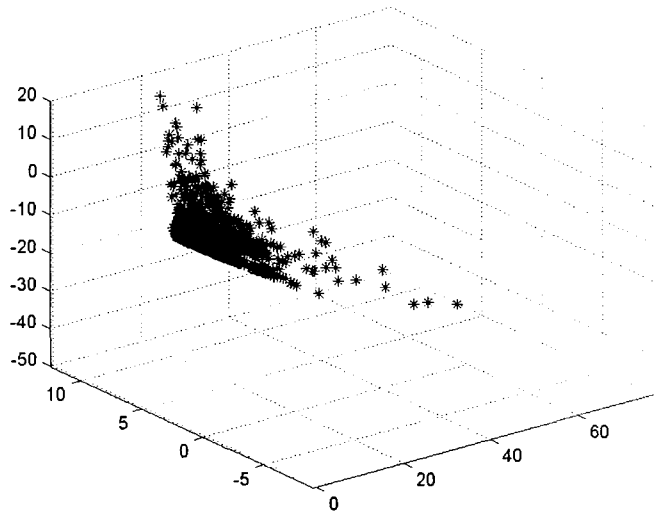


FIG. 5.5 – Répartition des 32790 événements dans l'espace engendré par les 3 premières composantes principales.

Il est aussi intéressant de voir comment les données peuvent être analysées grâce à l'ACP. En effet, dans un ensemble d'individus (un nuage de points) on peut définir des populations respectant certains critères. L'objectif du volet descriptif de l'analyse en composantes principales est d'étudier ces populations, de les caractériser et de rechercher les éventuelles différences à l'aide des nouvelles variables.

Nous allons diviser cette section en deux parties, la première dédiée aux séquences de phonoentérogramme et la deuxième aux sons abdominaux individuels.

Séquences de phonoentérogramme

Le nuage de points représentant les 3024 séquences de phonoentérogramme est représenté figure 5.3. A l'intérieur de ce nuage, on peut regrouper les points selon différents critères : les points issus d'un même patient, les points issus d'une des six voies, les points issus d'un regroupement temporel (on définit un regroupement temporel comme une suite de séquences). On considère que chaque population peut être représentée par un individu moyen, égal au barycentre du nuage des points de cette population. Les positions (proches ou éloignées) de ces barycentres dans l'espace des composantes principales peuvent indiquer les éventuelles différences ou similarités entre les populations. Dans le cadre de l'ACP, l'approche traditionnelle étudie les projections de ces barycentres sur les plans principaux (définis par les axes des composantes principales).

Nous analysons les différents groupes de séquences de phonoentérogramme selon cette méthode. Les 3024 séquences peuvent être réparties en groupes de plusieurs manières :

1. **Les groupes $\mathbf{v_1}$ à $\mathbf{v_6}$**, définis comme les ensembles constitués des séquences enregistrées par chacune des six voies (donc à des localisations spatiales différentes). Ce regroupement ne tient compte ni du numéro de la séquence dans la succession des 168 minutes d'acquisition, ni du patient. On obtient donc six groupes, chacun contenant $168 \cdot 3 = 504$ séquences d'une minute. L'objectif est la classification des voies en fonction de leurs caractéristiques.
2. **Les groupes $\mathbf{s_1}$ à $\mathbf{s_{12}}$**, définis comme les ensembles de séquences enregistrées par tranches de 14 minutes (s_1 regroupe les séquences enregistrées de la minute 1 à la minute 14, s_2 de 15 à 29, etc., jusqu'à s_{12}, de 155 à 168). Ce regroupement ne tient compte ni de la voie d'acquisition, ni du patient. Les douze groupes contiennent chacun $14 \cdot 3 \cdot 6 = 252$ points. L'objectif est l'étude de la séquence temporelle – les phases successives du cycle digestif peuvent-elles être différenciées selon leurs caractéristiques ?
3. **Les groupes $\mathbf{p_1}$ à $\mathbf{p_3}$**, contenant l'ensemble des séquences par patient. L'effectif de chacun de ces trois groupes est de $168 \cdot 6 = 1008$ séquences. L'objectif est de différencier les patients.

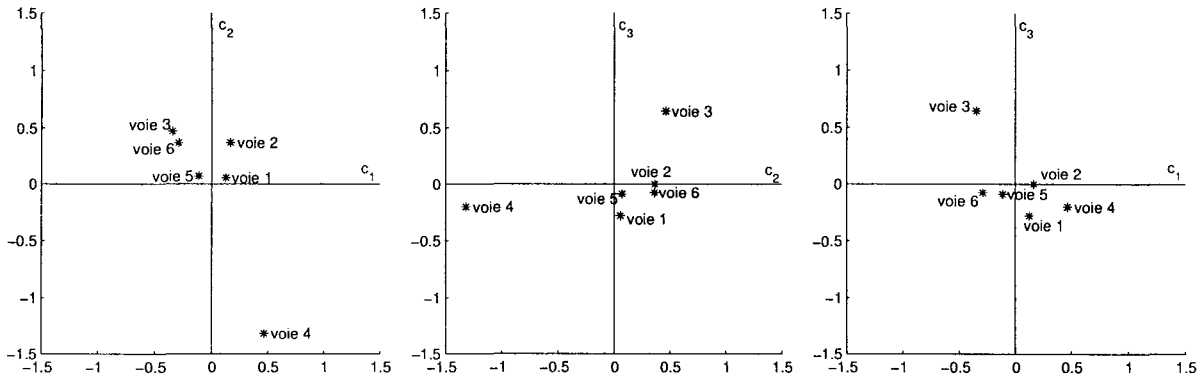
A l'intérieur de chaque groupe on peut définir des sous-groupes : par exemple, on peut étudier la différenciation des voies pour chaque patient, ou encore l'évolution temporelle pour chaque voie ou chaque patient. Nous présentons ici quelques résultats que nous estimons intéressants, sachant toutefois que nos considérations ont une valeur limitée, car elles doivent être validées sur un nombre plus important de patients et à l'aide d'un expert[50]. Cependant, l'objectif que nous nous sommes fixé est le développement d'un outil d'analyse des phonoentérogrammes, et les exemples que nous présentons ici montrent que la méthode proposée est capable de différencier plusieurs états physiologiques, sans pour autant pouvoir associer une interprétation médicale.

Différenciation des voies. Les résultats de la projection des barycentres des groupes v_1 à v_6 sur les plans définis par les axes principaux sont présentés figure 5.6. Un premier constat porte sur la différence que la composante principale c_2 révèle entre la voie v_4 et les autres voies.

Selon l'interprétation que nous avons donnée aux composantes principales dans la section 5.1.1, la voie v_4, située dans la zone sous-ombilicale et affectée principalement par l'activité de l'intestin grêle (voir figure 2.4, page 26), est en moyenne la plus riche en événements parmi les six (alors que v_3 et v_6 semblent les plus pauvres). La différence entre v_4 et v_3 se confirme visuellement à l'inspection du nuage de points (figure 5.7).

On peut également constater que la voie v_3 est la seule à être notablement supérieure à la moyenne selon c_3 (composante liée à la fréquence moyenne des sons) : les événements apparaissant

[50]Un protocole d'enregistrement et d'étude de phonoentérogrammes pathologiques proposé par le Professeur F. Guillemin est envisagé en collaboration avec le Centre Alexis Vautrin de Nancy.



(a) plan c_1–c_2 (b) plan c_2–c_3 (c) plan c_1–c_3

FIG. 5.6 – Projections des voies v_1 à v_6 sur les plans principaux.

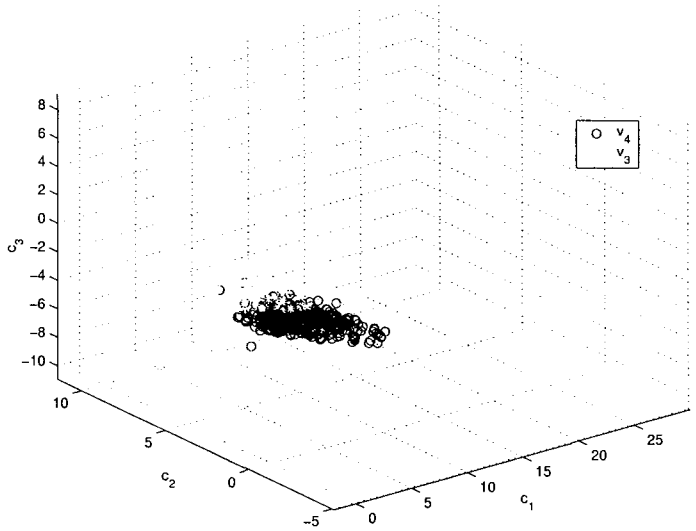


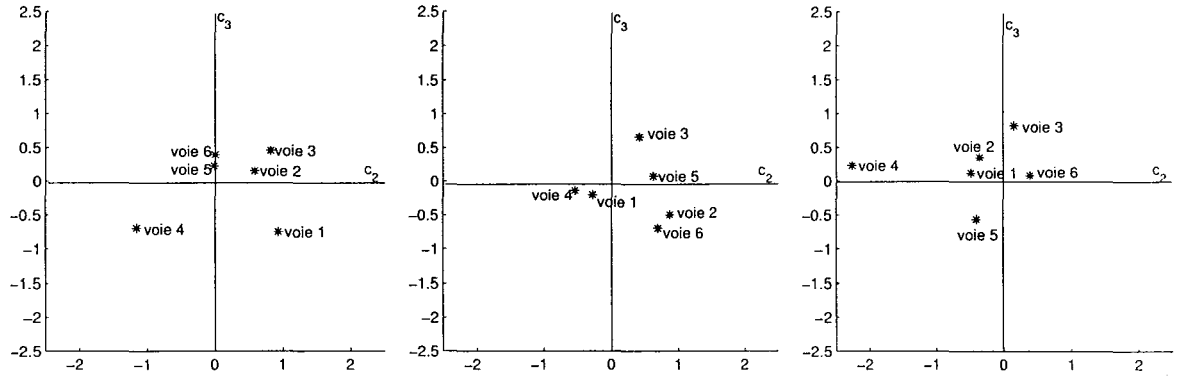
FIG. 5.7 – Répartition 3D des séquences des voies v_3 (504 points) et v_4 (504 points) dans l'espace engendré par les 3 premières composantes principales.

à cet endroit de l'abdomen (cadran inférieur gauche, début du côlon ascendant) semblent en général de plus haute fréquence que les autres.

Les remarques précédentes se confirment lors de l'étude détaillée patient par patient : le plan principal $c_2 - c_3$, qui est le plus adapté à la visualisation des différences signalées ci-dessus, est présenté séparément pour les trois patients sur la figure 5.8. La voie v_4 est toujours la plus riche en événements, avec un écart encore plus grand par rapport aux autres pour le patient p_3.

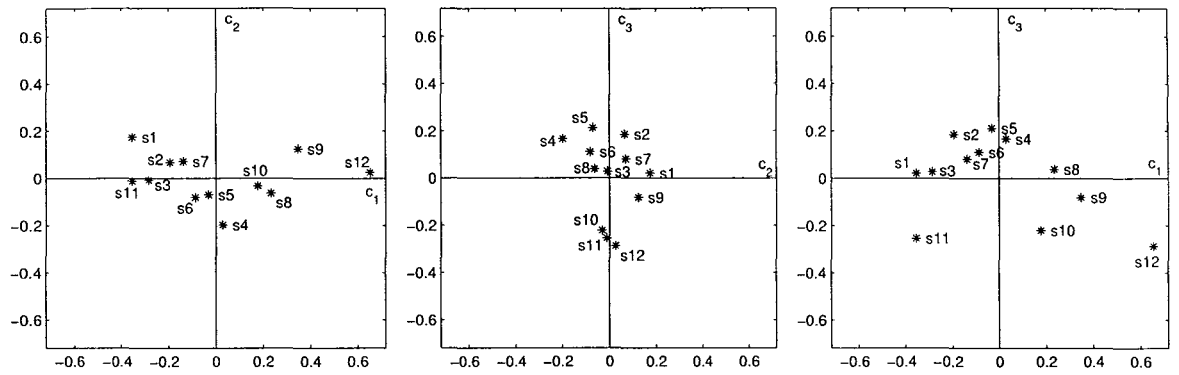
Différenciation des séquences. Les résultats de la projection des barycentres des groupes de séquences s_1 à s_{12} sur les plans définis par les axes principaux sont présentés figure 5.9.

Les différences sont moins nettes qu'entre les voies. En considérant le troisième axe principal c_3, on observe que les valeurs diminuent vers la fin du cycle digestif et deviennent négatives : l'activité sonore en période postprandiale semble être de plus haute fréquence que celle de la



(a) patient p_1 (b) patient p_2 (c) patient p_3

FIG. 5.8 – Projections des voies v_1 à v_6 sur les plans $c_2 - c_3$ pour les trois patients.



(a) plan c_1–c_2 (b) plan c_2–c_3 (c) plan c_1–c_3

FIG. 5.9 – Projections des séquences s_1 à s_{12} sur les plans principaux.

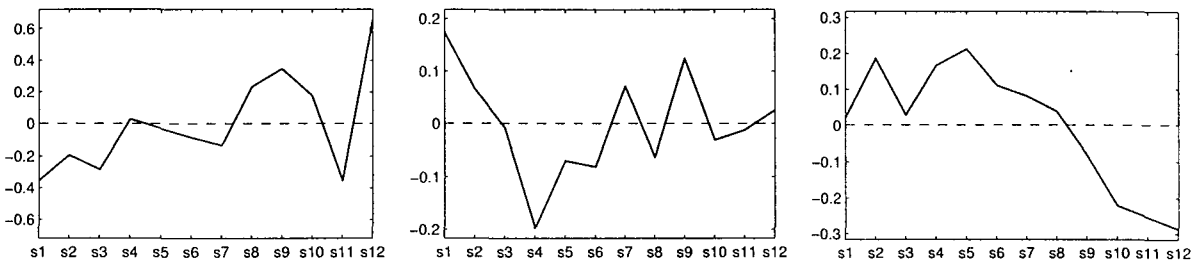
période suivante (jusqu'à la fin des 168 minutes d'enregistrement). On peut noter également que les valeurs moyennes de la première composante principale ont une tendance croissante au cours du temps (voir figure 5.10, qui montre l'évolution des séquences selon les axes principaux).

Certes, cette interprétation nécessite d'être confirmée car elle n'est pas évidente quand on considère les différents patients : par exemple, la diminution de la fréquence moyenne c_3 au cours du temps reste vraie pour les trois patients, mais son évolution et le moment du passage en dessous de la moyenne sont différents (périodes postprandiales ayant des caractéristiques variables en fonction du patient ?) (figure 5.11). De plus, les évolutions ne sont pas identiques sur toutes les voies, et sur deux voies (v_2 et v_4) la fréquence moyenne en fin d'enregistrement est légèrement supérieure à la moyenne.

Différenciation des patients. Les projections des barycentres des nuages de points correspondant à chaque patient sont présentées figure 5.12.

Le patient p_3 semble se détacher des deux autres selon l'axe c_2, ce qui indique un nombre plus

important d'événements. D'ailleurs, cette distinction était déjà visible figure 5.8, qui montre de plus que c'est la voie v_4 qui en est responsable. Cependant, il n'y a pas de différence significative



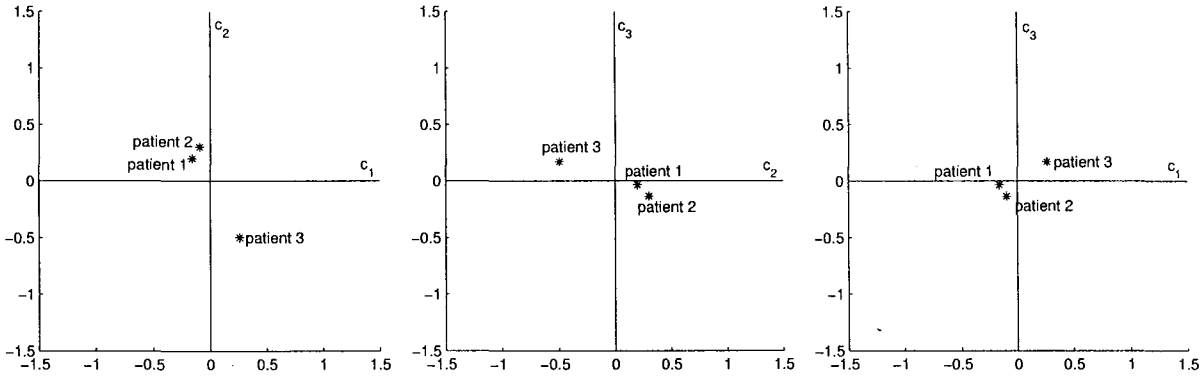
(a) axe c_1 (b) axe c_2 (c) axe c_3

FIG. 5.10 – Évolution temporelle des séquences selon les axes principaux. $c_1 \approx$ niveau sonore, $c_2 \approx$ richesse/pauvreté en événements, $c_3 \approx$ fréquence moyenne.



(a) patient p_1 (b) patient p_2 (c) patient p_3

FIG. 5.11 – Évolution des séquences s_1-s_{12} selon l'axe c_3 pour les trois patients. Les valeurs négatives doivent être interprétées comme inférieures à la moyenne générale, sur l'ensemble des points. Les lignes en pointillés représentent les moyennes par patient.



(a) plan c_1-c_2 (b) plan c_2-c_3 (c) plan c_1-c_3

FIG. 5.12 – Projections des patients p_1 à p_3 sur les plans principaux.

d'évolution temporelle dévoilée par l'étude des séquences successives. On peut donc conclure que la différence entre le patient p_3 et les autres est due à une voie v_4 plus riche en sons tout au long des 168 minutes d'enregistrement. Cette interprétation est confirmée par la comparaison des évolutions de l'indice c_2, calculée sur la voie v_4 pour les trois patients (figure 5.13).

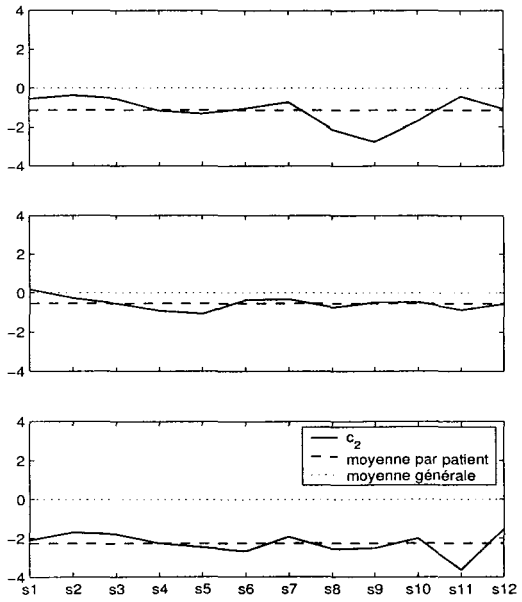


FIG. 5.13 – Comparaison des évolutions de c_2 pour la voie v_4, pour les trois patients. On observe que cette voie est en général en dessous de la moyenne générale (la voie est plus riche, voir figure 5.6). Le patient p_3 se différencie des autres grâce à une richesse en événements encore plus importante.

Cas pathologique. En dehors du protocole de mesure défini (section 2.2.1), nous disposons d'un enregistrement pathologique (gastroentérite). La durée d'enregistrement a été beaucoup moins importante (7 minutes) et l'alimentation n'a pas été standardisée, comme pour les autres phonoentérogrammes que nous avons traités. Après pré-traitement et extraction de caractéristiques, nous avons obtenu un groupe de 42 séquences élémentaires de phonoentérogramme pathologique et les indices d'activité associés.

Il y a deux solutions pour placer ces points dans le même espace de représentation que les séquences que nous avons étudiées jusqu'à présent :

- soit on les intègre aux données initiales, puis on re-normalise et on re-calcule l'espace de représentation et les composantes principales du nouveau nuage de points ;
- soit on garde l'espace défini antérieurement et on calcule les coordonnées des nouveaux points dans cet espace (on les normalise par rapport aux données d'origine et on les projette sur les axes principaux).

Nous avons adopté la deuxième solution. D'abord, parce que le nombre de nouveaux points est très faible par rapport aux points ayant servi à l'ACP initiale et donc leur influence est faible.

Ensuite, du point de vue méthodologique, il nous a semblé intéressant de garder l'espace d'origine, calculé à partir d'un ensemble de phonoentérogrammes normaux et donc *a priori* facile à obtenir et indépendant de toute éventuelle pathologie recherchée. Le risque inhérent de cette approche est lié au nombre d'axes principaux gardés pour la représentation : on rappelle que les axes de l'espace ont une importance décroissante et que les trois premiers préservent l'essentiel de la variance du nuage de points non pathologiques. Cependant, il n'est pas garanti que le nuage

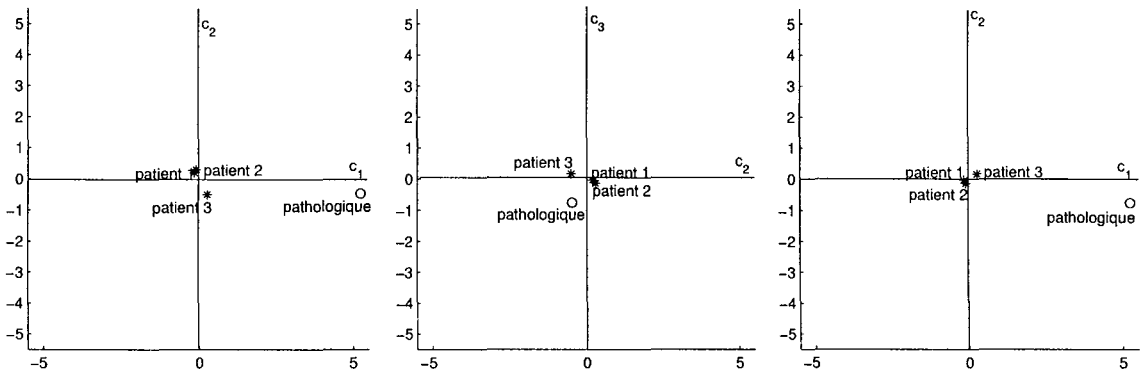
pathologique soit bien préservé dans cet espace réduit. En l'occurrence, les pourcentages de la dispersion du nuage pathologique expliqués par les axes principaux établis antérieurement sont différents de ceux du nuage initial (tableau 5.7, par rapport au tableau 5.2) et ce nuage conserve seulement 69,1% de la variance après la re-projection sur l'espace tridimensionnel engendré[51].

	var	%	% cumulé
c_1	47,19	52,3	52,3
c_2	10,25	11,4	63,7
c_3	4,95	5,5	69,2
c_4	12,44	13,8	83,0
c_5	6,12	6,8	89,8
c_6	1,40	1,5	91,3
c_7	6,37	7,1	98,4
c_8	0,67	0,7	99,1
c_9	0,81	0,9	100,0

TAB. 5.7 – Variances des composantes principales du signal pathologique (dans l'espace défini à partir des phonoentérogrammes normaux) et leurs contributions, en pour-cent, à la variance des données du patient malade.

Vu le faible effectif du groupe de séquences pathologiques, nous ne l'avons pas séparé en sous-groupes. Par conséquent, il est représenté par un nuage unique de 42 points. La projection de son barycentre sur les plans principaux permet une comparaison entre ce phonoentérogramme et les trois phonoentérogrammes des patients sains (figure 5.14), et on remarque que :

- le niveau sonore c_1 du patient malade est beaucoup plus élevé ;
- la "fréquence moyenne" c_2 est légèrement plus basse ;
- la richesse en événements c_2 n'est pas discriminante entre ce signal et celui du patient p_3.



(a) plan c_1–c_2 (b) plan c_2–c_3 (c) plan c_1–c_3

FIG. 5.14 – Projections des patients sains p_1–p_3 et du patient malade sur les plans principaux. L'échelle est modifiée par rapport à la figure 5.12 pour prendre en compte les coordonnées du phonoentérogramme pathologique.

[51] Ce même problème se pose implicitement pour tout sous-groupe du nuage initial de points – il ne sera pas forcément bien représenté dans un espace construit pour l'ensemble.

Sons abdominaux individuels

L'analyse des sons abdominaux individuels peut se faire selon la même méthodologie que celle des séquences de phonoentérogramme. Les groupes étudiés et comparés sont constitués des sons individuels enregistrés voie par voie, durant des séquences successives de 15 minutes ou patient par patient.

Cependant, il y a deux inconvénients importants à cette approche :

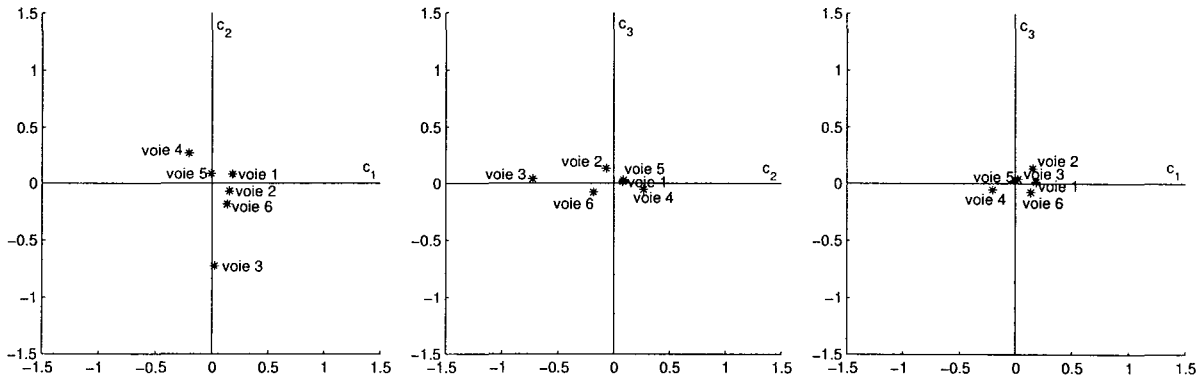
1. Comme l'interprétation de l'analyse en composantes principales est fondée sur la comparaison des moyennes des différents groupes, elle est biaisée par les effectifs de ces groupes. Dans le cas des séquences de phonoentérogramme, l'effectif des groupes comparés était identique par définition, alors qu'ici il n'y a pas de raison de supposer que le même nombre de sons sera produit à un endroit ou un autre de l'abdomen (comparaison des voies), au début et à la fin du cycle digestif (comparaison des séquences) ou par différentes personnes (comparaison des patients). D'ailleurs, cette hypothèse est fausse : l'analyse en composantes principales des séquences présentée dans la section précédente compare les groupes, entre autres, en fonction de leur richesse en événements, et elle montre par exemple que le patient p_3 produit plus de sons que p_1 et p_2. Si les différences entre les effectifs et entre les caractéristiques moyennes des patients sont importantes, les comparaisons globales entre les voies et les séquences ne peuvent plus être interprétées, car elle seront en fait dues aux différences au sein du groupe de sons produits par le patient le plus actif.

2. Comme l'espace de représentation des sons retenu après l'ACP est engendré par des axes qui ont une signification proche des axes de l'espace utilisé pour la représentation des séquences, l'information risque d'être redondante. La raison est simple : le barycentre d'un groupe de sons est un son "représentatif" caractérisé par un niveau sonore moyen (c_1 dans l'espace des sons individuels), une fréquence moyenne (c_2) et une durée moyenne (c_3). Or l'axe de la fréquence moyenne et celui du niveau sonore sont des axes proches de ceux de l'espace des séquences, et un événement moyen d'un patient ne sera pas loin de la séquence moyenne du même patient, vu qu'elle est calculée à son tour à partir des sons qui la composent.

   Pour illustrer ces propos, il est suffisant de comparer les figures 5.15 et 5.6 : la voie v_3 se détache (faiblement) des autres par sa fréquence plus élevée (c_3 dans l'espace des séquences, c_2 dans l'espace des sons individuels) dans les deux cas. Dans l'espace des sons, on perd l'information sur la richesse en événements de la voie v_4. En revanche, la figure 5.15 suggère que les événements issus de cette voie sont plutôt basse-fréquence, ce qui n'était pas visible sur la figure 5.6.

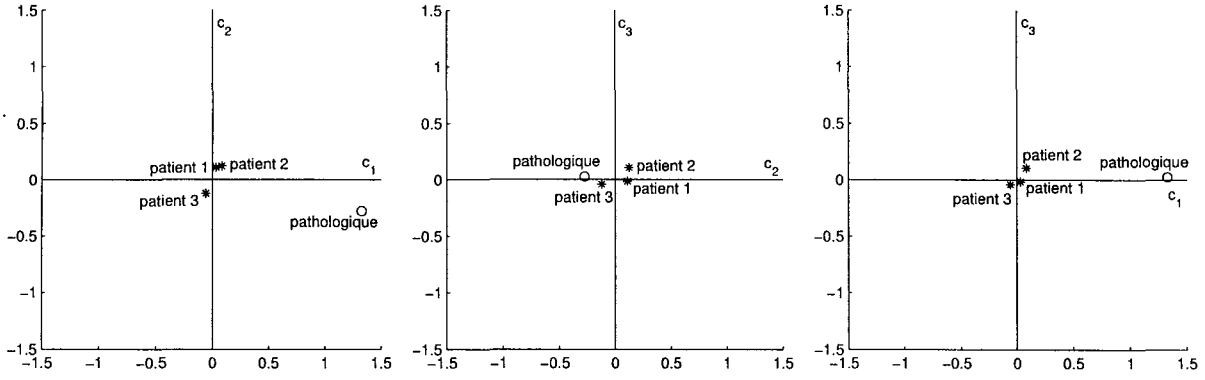
La même redondance de l'information est constatée pour le cas pathologique (figure 5.16) : le patient malade se détache des autres par un niveau sonore plus important et, légèrement, par une fréquence plus basse, ce qui était déjà visible sur la figure 5.14.

A la vue des constats précédents, l'analyse en composantes principales appliquée comme un outil de classification (de différenciation entre les groupes) dans l'espace des sons individuels n'apporte pas de nouvelles informations, elle présente une certaine redondance par rapport à l'ACP dans l'espace des séquences. De plus, son interprétation risque d'être biaisée par la distribution des sons dans les groupes (classes). Un choix entre les deux espaces n'est pas possible pour l'instant, car nous ne disposons pas d'une base de signaux normaux et pathologiques suffisante. D'ailleurs, une étude sur davantage de cas pourrait confirmer l'utilité des deux représentations, s'il s'avère qu'elles sont sur certains points complémentaires.



(a) plan c_1–c_2 (b) plan c_2–c_3 (c) plan c_1–c_3

FIG. 5.15 – Projections des voies v_1 à v_6 sur les plans principaux de l'espace des sons individuels.



(a) plan c_1–c_2 (b) plan c_2–c_3 (c) plan c_1–c_3

FIG. 5.16 – Projections des sons individuels normaux des patients p_1–p_3 et des sons pathologiques sur les plans principaux.

5.2 Classification non-supervisée

Comme nous l'avons déjà remarqué, l'analyse en composantes principales ne constitue pas un outil de classification. En effet, elle estime plutôt le rôle des composantes principales dans la différenciation des groupes (classes), et elle ne cherche pas à mieux différencier ces groupes. Les résultats obtenus sont néanmoins intéressants pour notre problème, son application est simple et l'interprétation est relativement facile. Elle n'est cependant pas capable de trouver des groupes et on ne peut pas détecter, par exemple, un type particulier d'activité abdominale ou un son particulier s'il n'est pas, d'une part, spécifique à un groupe prédéfini, et d'autre part, suffisamment bien représenté par la moyenne de ce groupe. Autrement dit, on n'arrivera jamais à détecter un type d'activité apparaissant durant seulement quelques minutes du cycle digestif sur plusieurs voies, ou encore un événement rare comme un son spécifique à une pathologie.

Le problème que l'on se pose est donc : existe-t-il des classes de séquences ou de sons qui soient définies "naturellement" grâce à leurs variables explicatives (niveau sonore, fréquence moyenne,

etc.) et non pas arbitrairement, par rapport à leur lieu d'émission ou instant d'apparition ? Comment peut-on les trouver ? Notre objectif immédiat est donc la découverte des classes naturelles, si elles existent, dans l'ensemble des points (séquences ou sons individuels). Le deuxième objectif est l'utilisation de ces classes comme étalon : tout nouveau phonoentérogramme sera traité et comparé à cet ensemble d'apprentissage, afin d'apprécier sa normalité ou les éventuelles perturbations physiologiques.

Comme les connaissances sur la classification de l'activité sonore abdominale sont quasiment toujours subjectives et non formalisées, la démarche suivie a été la classification non-supervisée.

Cette section commence par quelques considérations générales et un bref état de l'art sur les méthodes de classification (reconnaissance des formes) non-supervisée. Cette présentation n'a pas l'ambition d'être exhaustive, car les travaux sont extrêmement nombreux et les algorithmes sont souvent créés sur mesure pour les différentes applications (ce qui est incontournable dans le domaine).

Ensuite, nous étudions leur applicabilité dans le domaine de la phonoentérographie, nous présentons les hypothèses de travail et nous proposons un algorithme de classification non-supervisée. La dernière partie présente et commente les résultats.

5.2.1 Méthodes de classification non-supervisée

Comme les connaissances dans ce domaine sont subjectives et peu structurées, nous employons une méthode de classification non-supervisée afin de construire l'ensemble d'apprentissage. Cette méthode doit permettre de trouver des classes "naturelles" homogènes dans l'ensemble des points de l'espace, sans connaître leur nombre, leurs positions, leurs tailles, leurs formes et leurs effectifs.

Il est nécessaire de préciser ce que l'on entend par la notion de "classe naturelle". Cette notion est liée à celle de similarité : une classe "naturelle" contient des éléments similaires entre eux et différents des éléments appartenant à d'autres classes. En représentant les classes comme des nuages de points dans l'espace des caractéristiques, la plus évidente mesure de similarité/dissimilarité entre deux éléments (points) est la distance qui les sépare. Il est donc nécessaire de définir la métrique utilisée, car elle influence d'une manière essentielle la classification.

Le problème énoncé au premier paragraphe de cette section est beaucoup trop général : il y a trop de paramètres inconnus (la métrique, le nombre de classes, leurs positions, leurs tailles, leurs formes et leurs effectifs) pour qu'il puisse être résolu. D'ailleurs, dans [40], les auteurs montrent qu'il est impossible de trouver LE meilleur algorithme de classification, car les performances dépendent toujours du problème (*No Free Lunch Theorem*) et de l'espace de caractéristiques dans lequel on travaille (*Ugly Duckling Theorem*). Nous présentons dans la suite brièvement les métriques les plus courantes et les principales familles d'algorithmes de classification. Ensuite, nous proposons des hypothèses de travail restrictives, afin de limiter les choix possibles.

Mesures de distance. Vu la construction de notre espace de caractéristiques, nous allons nous limiter à présenter les métriques les plus courantes associées aux données numériques. En fait, ces métriques sont dérivées de la distance générale de Minkowski. Entre deux points x_i et x_j elle se définit comme :

$$d(x_i, x_j) = ||x_i - x_j||_p,$$

où $||\cdot||_p$ est la norme usuelle dans L_p. La distance la plus utilisée est la **distance euclidienne**, obtenue pour $p = 2$, mais la distance *Manhattan* (p=1) est aussi courante. Les distances de type Minkowski présupposent un espace engendré par des axes (variables explicatives) orthonormés.

Une autre métrique courante est celle de Mahalanobis, qui prend en compte la matrice de variance-covariance Σ des variables explicatives et compense la non-orthogonalité des axes :

$$d_M(x_i, x_j) = \sqrt{(x_i - x_j)^T \Sigma^{-1} (x_i - x_j)}. \tag{5.1}$$

Algorithmes non-supervisés. Un passage en revue rapide de la littérature nous révèle l'existence d'une pléthore d'algorithmes de classification non-supervisée adaptés à des situations plus ou moins spécifiques. Plusieurs travaux de synthèse [10, 42, 59, 64] proposent différentes classifications de ces algorithmes, suivies par une description de leur principe, de leurs avantages et inconvénients. Les frontières entre les familles algorithmiques sont loin d'être nettes, et de nombreuses méthodes partagent des caractéristiques communes à plusieurs familles [10, 64].

Selon [10, 42, 59, 79, 100], on peut distinguer deux grandes familles de méthodes de classification non-supervisée (*clustering*) – les méthodes de partitionnement et les méthodes hiérarchiques[52]. Nous allons présenter dans la suite ces deux familles de méthodes, en insistant sur les méthodes hiérarchiques que nous allons utiliser en priorité dans le cas des sons abdominaux.

Les méthodes de partitionnement cherchent à trouver un nombre k de régions homogènes dans l'espace de classification. En règle générale, ce nombre doit être connu avant l'application de ces méthodes.

Ces méthodes, qui peuvent souvent être interprétées géométriquement, cherchent à estimer la position des k classes à l'aide de points considérés représentatifs (par exemple les barycentres, dans l'algorithme *k-means*) et à attribuer chaque point au barycentre le plus proche. La recherche des points représentatifs et l'allocation des points sont itératives : l'initialisation est aléatoire et l'algorithme converge quand une partition stable est atteinte. L'avantage principal de ces algorithmes est leur simplicité. Cependant, ils dépendent de l'initialisation et ils favorisent fortement la création de classes de forme sphérique de même rayon.

D'un point de vue probabiliste, la répartition des points dans l'espace de classification peut être vue comme issue d'un mélange de densités de probabilité, qui doivent être estimées. Les méthodes paramétriques modélisent ces densités par un jeu de paramètres : typiquement, on considère des lois de densités de probabilité gaussiennes (forme connue) dont on estime les moyennes (les positions des barycentres dans l'algorithme *k-means*), les variances (les tailles) et les probabilités *a priori* (l'effectif). L'estimation de ces paramètres est en général réalisée par des méthodes de maximum de vraisemblance. Comme montré par Duda, Hart et Stork [40], le problème d'estimation des paramètres des mélanges gaussiens complètement inconnus et sans contraintes n'accepte pas une solution globale optimale, car le maximum de vraisemblance global tend vers l'infini. En pratique, on cherche une solution sous contraintes à l'aide d'un algorithme itératif qui converge vers un maximum local de la vraisemblance, comme les algorithmes EM (*Expectation-Maximisation*) adaptés à la classification [10, 20, 42, 45, 59].

L'avantage de ces méthodes est la possibilité qu'elles offrent de trouver des classes ellipsoïdales de tailles et effectifs différents. L'inconvénient est lié à la nature parfois fortement sous-optimale des solutions trouvées (car les algorithmes sont très sensibles à l'initialisation mais aussi, comme pour les méthodes géométriques, dans l'imposition du nombre de classes).

Des méthodes plus générales proposent des estimations non-paramétriques des lois de probabilité du mélange, par des fonctions de noyau comme les fenêtres de Parzen [40], mais elles deviennent rapidement inapplicables dès que la dimension de l'espace augmente et l'estimation perd sa fiabilité.

[52] Kolatch [64] distingue en plus les méthodes locales, qui construisent les classes itérativement en suivant la connectivité des régions denses dans l'espace.

Enfin, les méthodes locales, qui définissent les classés simplement comme des régions denses, estiment cette densité en comptant le nombre de points à l'intérieur des régions connexes de l'espace. Dans certains cas (classes de formes irrégulières), ces méthodes peuvent fournir de meilleurs résultats.

Comme nous l'avons précisé précédemment, le nombre de classes k doit être spécifié avant l'application des méthodes de partitionnement et il est en règle générale fixé par l'utilisateur, en fonction du problème réel à résoudre. Dans la littérature on trouve aussi des méthodes permettant d'estimer k à l'aide de différents critères de "bonne classification". Nous rappelons ici seulement leurs principes (voir [10, 109] pour plus de détails), qui dépendent du cadre du problème de classification, géométrique ou probabiliste :

- les critères géométriques les plus usuels calculent le coefficient de séparation entre deux classes i et j de deux manières : soit à partir de leurs paramètres (les barycentres-moyennes μ_i et μ_j et les rayons-écarts-types σ_i et σ_j) comme l'index Davies-Bouldin, soit directement à partir des distances entre les points d'une même classe et ceux des autres classes (coefficient de Dunn ou coefficient Silhouette) ;
- dans le cas probabiliste, une bonne classification est donnée par le compromis entre l'adéquation du modèle aux données (au sens du maximum de vraisemblance) et le nombre de paramètres à estimer (moyennes, variances, probabilités *a priori*). Les critères les plus connus sont MDL (*Minimum Description Length*), BIC (*Bayesian Information Criterion*), AWE (*Approximate Weight of Evidence*), etc. [10, 42, 45].

L'utilisation d'une de ces méthodes revient toujours à la résolution successive de plusieurs problèmes de classification pour différents k, suivie du choix de la meilleure solution selon le critère de validation imposé.

Les méthodes hiérarchiques ne construisent pas une partition de l'espace en plusieurs classes et proposent une succession de partitions emboîtées.

Ces méthodes peuvent être divisives (descendantes) ou agglomératives (ascendantes). Les méthodes descendantes commencent par une "supra-classe" unique, qui est divisée ensuite successivement jusqu'au niveau des individus. Cependant, en pratique seules les méthodes ascendantes sont utilisées : les individus proches sont groupés en classes, qui sont à leur tour groupées en "supra-classes" jusqu'à l'obtention d'une classe unique contenant tous les points.

Il existe plusieurs manières d'associer les classes proches en supra-classes [1, 10, 40, 100] ou, autrement dit, de définir ce que l'on entend par la distance $\mathcal{D}_{ij}$ entre deux classes i et j (x_i et x_j sont des points appartenant respectivement aux classes i et j) :

- la distance du plus proche voisin (*single linkage*)

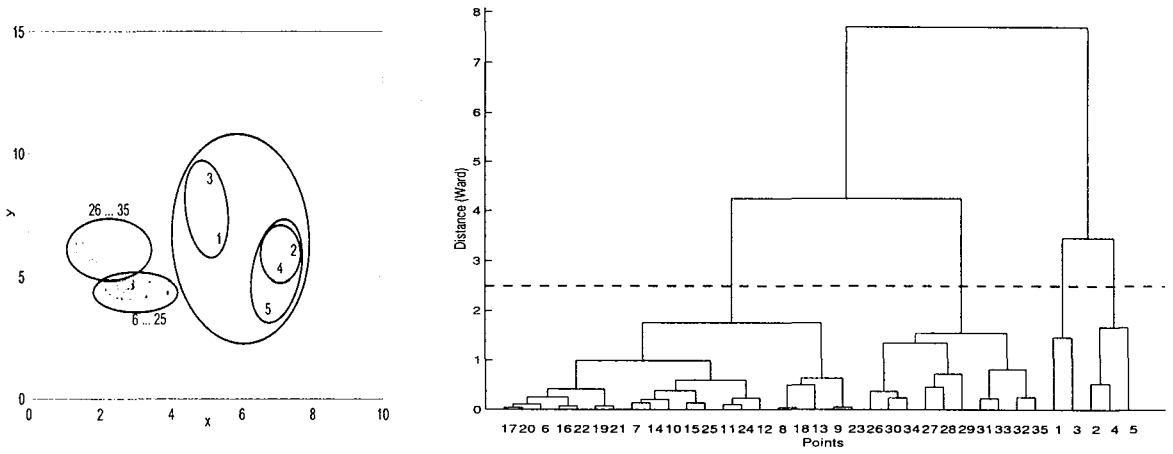
$$\mathcal{D}_{ij,min} = \min_{i,j} d(x_i, x_j),$$

  considère que la distance entre deux classes est donnée par la distance minimale qui les sépare ;
- la distance du plus éloigné voisin (*complete linkage*)

$$\mathcal{D}_{ij,max} = \max_{i,j} d(x_i, x_j)$$

  considère que la distance entre deux classes est donnée par la distance entre les deux points les plus éloignés appartenant chacun à une des classes ;
- la distance moyenne (*average linkage*)

$$\mathcal{D}_{ij,avg} = \frac{1}{n_i n_j} \sum_{i,j} d(x_i, x_j)$$



(a) nuages de points

(b) dendrogramme correspondant

FIG. 5.17 – Exemple de classification hiérarchique. Les classes emboîtées se retrouvent dans la structure arborescente du dendrogramme. La ligne en pointillé représente une possibilité de découpage du dendrogramme, qui montre les inconvénients du découpage à un niveau fixe : les points 1... 5 sont séparés en deux classes.

calcule la distance entre deux classes comme la moyenne des distances point par point ;

- la distance des barycentres (*mean linkage*)

$$\mathcal{D}_{ij,\mu} = d(\mu_i, \mu_j)$$

calcule la distance entre deux classes comme la distance entre les barycentres ;

- la distance de Ward

$$\mathcal{D}_{ij,ward} = \frac{n_i n_j}{n_i + n_j} \mathcal{D}^2_{ij,\mu}$$

propose l'association des classes qui diminue le moins possible l'inertie interclasse[53].

La structure emboîtée des classes peut être vue comme une structure arborescente, représentée à l'aide des dendrogrammes (figure 5.17). Les points situés aux extrémités des branches de l'arbre (les feuilles) sont les éléments à classifier, et la hauteur de la branche qui les relie est donnée par la distance entre les points. Le même raisonnement est mené pour relier deux classes – la hauteur de la branche est égale à la distance qui les sépare (définie selon les expressions ci-dessus).

Cette approche n'aboutit pas à une solution directe du problème de classification : le dendrogramme fournit une arborescence hiérarchique de classes et non pas un partitionnement. Ce dernier est obtenu par découpage – il faut tailler l'arbre afin d'obtenir les branches représentant les classes, ce qui pose un nouveau problème : où découper ?

Aucune des méthodes d'association listées ci-dessus n'impose la spécification de la forme, de la position ou de l'effectif des classes, ce qui est un avantage certain quand les connaissances *a priori* sont faibles. Cependant, cette généralité soulève des difficultés supplémentaires : quelle méthode choisir ou, autrement dit, comment choisir le meilleur dendrogramme ? Cette question en sous-entend une autre : comment interpréter un dendrogramme ?

Les deux questions précédentes (comment interpréter un dendrogramme et comment le découper) sont reliées au problème fondamental de la validité de toute méthode de classification,

[53] L'inertie interclasse se définit comme la variance du nuage des barycentres, chaque barycentre ayant un poids égal à l'effectif de la classe correspondante [100].

et à l'impossibilité de principe de choisir LA meilleure méthode (*No Free Lunch Theorem*). Cependant, comme dans le cas des méthodes de partitionnement (indices Davies-Bouldin, Dunn, etc.), dans la littérature on trouve plusieurs manières d'apprécier la validité d'une classification hiérarchique (et donc la lecture d'un dendrogramme).

Les deux indices les plus utilisés sont le coefficient de corrélation *cophenetique* [1, 19] et le coefficient d'inconsistance [1] :

- le **coefficient cophenetique** calcule la corrélation des distances entre les points avec les distances entre les classes d'affectation de ces points. Plus précisément, pour toutes les paires de points (x_i, x_j), x_i appartenant à la classe i et x_j appartenant à la classe j, on peut calculer la distance $d_{ij}(x_i, x_j)$ (entre ces deux points, selon la métrique choisie) et la distance $\mathcal{D}_{ij}$ (entre les deux classes, selon la méthode d'association choisie). Le coefficient cophenetique est la corrélation entre les vecteurs $\mathbf{d}$ et $\mathcal{D}$, calculée à chaque niveau du dendrogramme (chaque fois qu'on associe des sous-classes). Ce coefficient peut être utilisé pour mesurer la validité de la combinaison métrique-méthode d'association [1] ou pour décider du niveau de coupure du dendrogramme [19] : le regroupement des sous-classes s'arrête quand on constate une forte diminution du coefficient cophenetique.
  La réponse donnée par le coefficient cophenetique au problème du découpage a un point faible : il fixe un niveau unique pour toutes les branches, ce qui peut être gênant quand les classes ont des densités différentes (dans l'exemple présenté figure 5.17, le seuil fixé figuré par la ligne en pointillé sépare la classe la plus rare en deux classes distinctes). Quant à la validation de la méthode d'association choisie, elle n'a pas vraiment de sens avant le découpage, vu que le vecteur $\mathcal{D}$ des distances entre les classes dépend de celui-ci (on peut toutefois faire une moyenne des coefficients cophenetiques, comme proposé sous Matlab) ;
- le **coefficient d'inconsistance** ne cherche pas à apprécier la validité globale du dendrogramme, mais à détecter les sauts dans les longueurs des branches. D'une manière analogue aux critères de séparation pour les algorithmes de partitionnement, le coefficient d'inconsistance d'une branche compare sa longueur à la moyenne des longueurs des branches inférieures et normalise par rapport à leurs écarts-types. Quand les longueurs des branches varient d'une manière importante entre deux niveaux de l'arbre (la distance entre les classes reliées est beaucoup plus importante que les distances entre les sous-classes respectives), le coefficient d'inconsistance a une valeur importante qui peut être détectée par un seuillage adéquat.
  A la différence du découpage proposé par le coefficient cophenetique, celui obtenu grâce aux coefficients d'inconsistance est plus flexible : il permet des coupures à différents niveaux du dendrogramme, aux endroits signalés par les sauts de longueurs des branches. Il est nécessaire de remarquer qu'une branche qui relie deux feuilles (située donc sur le niveau le plus bas du dendrogramme) aura toujours par convention un coefficient d'inconsistance nul.

A notre avis, il est difficile d'utiliser les coefficients présentés ci-dessus pour choisir un algorithme hiérarchique spécifique, et il faut plutôt se rapporter au cadre applicatif. En pratique, on adapte la méthode d'association aux données et à la métrique de l'espace de classification. Par exemple, Saporta [100] recommande l'utilisation de la distance de Ward lorsqu'on travaille dans un espace muni d'une métrique euclidienne.

Enfin, concernant les méthodes de découpage, il faut choisir entre celles qui proposent un niveau fixe et celles qui s'adaptent à la structure locale du dendrogramme. La première méthode, la plus simple, consiste à fixer un seuil pour la longueur des branches, ce qui revient au découpage du dendrogramme à un niveau donné. Ce seuil peut être choisi en fonction du nombre de classes que l'on veut obtenir, si cette information est disponible ou, comme nous l'avons remarqué

précédemment, à l'aide du coefficient cophenetique.

La deuxième méthode utilise le coefficient d'inconsistance et respecte mieux, à notre avis, la répartition des données. Cependant, elle soulève le problème supplémentaire du choix du seuil : à partir de quelle valeur considère-t-on une branche comme inconsistante ?

5.2.2 Hypothèses de travail et algorithme

Nous devons examiner le problème concret de classification des séquences de phonoentérogramme et des sons abdominaux individuels afin de choisir une méthode de classification. Dans les deux cas, il s'agit de phénomènes physiques décrits par des variables explicatives ayant à leur tour une signification physique. Après l'analyse en composantes principales effectuée à l'étape précédente d'extraction de caractéristiques (section 5.1.1), les axes de nos espaces de classifications sont orthogonaux, ce qui implique que la matrice de variance-covariance Σ est diagonale. Les valeurs des éléments sur la diagonale sont différentes et correspondent aux variances selon chaque axe (composante principale). Afin de donner une importance similaire (du point de vue discriminant) aux axes, nous normalisons chaque variable explicative par son écart-type. Cette opération rend la matrice Σ unitaire, ce qui permet d'utiliser directement une **métrique euclidienne**.

Nous allons faire une hypothèse fondamentale, liée à la distribution des points à classifier dans l'espace : nous supposons qu'ils sont issus de plusieurs processus aléatoires de lois gaussiennes multidimensionnelles. Cette hypothèse est largement acceptée pour la modélisation des processus naturels et très souvent utilisée dans les algorithmes de classification. Dans un souci de généralité, nous n'imposons pas de restrictions supplémentaires sur les paramètres des classes gaussiennes ellipsoïdales que l'on recherche :

- leur nombre est inconnu ;
- leurs moyennes sont inconnues ;
- elles peuvent avoir des matrices de variance-covariance différentes et non-diagonales (ce qui implique des classes de formes et d'orientations différentes) ;
- elles peuvent avoir des probabilités *a priori* différentes et inconnues (ce qui implique des effectifs inconnus).

Selon les hypothèses de travail, nous nous plaçons dans un cadre probabiliste paramétrique dans lequel nous devons estimer les paramètres d'un mélange de plusieurs densités de probabilité. En plus du nombre de classes, la moyenne, la variance et les probabilités d'appartenance des points de chaque classe doivent être estimées par une méthode de maximum de vraisemblance. Les valeurs finales de ces paramètres sont obtenues après la convergence d'un algorithme itératif. Cependant, comme nous l'avons précisé précédemment, ceci constitue un problème d'optimisation délicat, difficile à résoudre sans une recherche quasi-exhaustive de toutes solutions possibles pour toutes les initialisations et leur comparaison selon un critère de bonne classification.

Nous proposons une autre approche, dont l'idée initiale est inspirée des méthodes de Fraley et Raftery [45] et le développement ultérieur s'apparente aux techniquès d'analyse discriminante [21, 40, 100]. L'idée générale est de trouver une solution sous-optimale initiale par un algorithme hiérarchique (comme proposé dans [45]) et d'améliorer et de stabiliser cette solution par rapport aux critères de classification fournis par l'analyse discriminante quadratique. Avant de décrire notre algorithme en détail, nous considérons nécessaire une présentation du principe de l'analyse discriminante quadratique.

Analyse discriminante

Cette méthode est classique en classification supervisée, une fois l'ensemble d'apprentissage construit. Elle suppose l'existence d'un ensemble de classes C_j ($j = 1 \ldots k$) défini *a priori* et cherche à attribuer un individu $\boldsymbol{x}$ (vu comme un vecteur de variables explicatives x_i, avec $i = 1 \ldots d$, d étant la dimension de l'espace de caractéristiques) à la classe la plus probable.

Pour un point $\boldsymbol{x}$, la probabilité d'appartenance à la classe C_j (c'est-à-dire la probabilité *a posteriori* de la classe C_j sachant $\boldsymbol{x}$) s'écrit (règle de Bayes) :

$$\begin{aligned} p(C_j|\boldsymbol{x}) &= \frac{p(\boldsymbol{x}|C_j)P(C_j)}{p(\boldsymbol{x})}, \text{ avec} \\ p(\boldsymbol{x}) &= \sum_{j=1}^{k} p(\boldsymbol{x}|C_j)P(C_j), \end{aligned} \tag{5.2}$$

où $p(\boldsymbol{x}|C_j)$ est la probabilité de $\boldsymbol{x}$ sachant C_j (c'est-à-dire la probabilité *a priori* d'appartenance à la classe C_j), $P(C_j)$ est la probabilité *a priori* de la classe C_j, et $p(\boldsymbol{x})$ la probabilité marginale sur l'ensemble des classes.

Décider de l'appartenance de $\boldsymbol{x}$ à une classe revient à choisir la classe C_j qui assure une probabilité *a posteriori* $p(C_j|\boldsymbol{x})$ maximale, ce qui revient à maximiser selon j la fonction $g(C_j)$:

$$g(C_j) = p(\boldsymbol{x}|C_j)P(C_j),$$

ou encore la fonction $l(C_j)$:

$$l(C_j) = \ln p(\boldsymbol{x}|C_j) + \ln P(C_j). \tag{5.3}$$

En considérant des classes gaussiennes C_j de moyennes $\boldsymbol{\mu}_j$ et matrices de variance-covariance $\boldsymbol{\Sigma}_j$, la probabilité *a priori* $p(\boldsymbol{x}|C_j)$ s'écrit :

$$p(\boldsymbol{x}|C_j) = \frac{1}{(2\pi)^{d/2}|\boldsymbol{\Sigma}_j|^{1/2}} \exp\left[-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu}_j)^T\boldsymbol{\Sigma}_j^{-1}(\boldsymbol{x}-\boldsymbol{\mu}_j)\right], \tag{5.4}$$

et l'appartenance à une classe se décide par la minimisation de :

$$\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu}_j)^T\boldsymbol{\Sigma}_j^{-1}(\boldsymbol{x}-\boldsymbol{\mu}_j) + \frac{1}{2}\ln|\boldsymbol{\Sigma}_j| - \ln P(C_j). \tag{5.5}$$

En pratique, on ne dispose pas des valeurs des paramètres $\boldsymbol{\mu}_j$ et $\boldsymbol{\Sigma}_j$ et on ne connaît pas la probabilité *a priori* $P(C_j)$. La solution courante est le recours aux estimations de maximum de vraisemblance, calculés à partir des points $\boldsymbol{z}_j$ ($j = 1 \ldots n_j$) qui constituent cette classe[54]. Ces estimations ont des expressions bien connues [20, 40, 100] :

$$\text{moyenne} \quad : \quad \hat{\boldsymbol{\mu}}_j = \frac{1}{n_j}\sum_{j=1}^{n_j} \boldsymbol{z}_j, \tag{5.6}$$

$$\text{matrice de variance} \quad : \quad \hat{\boldsymbol{\Sigma}}_j = \frac{1}{n_j}\sum_{j=1}^{n_j} (\boldsymbol{z}_j - \hat{\boldsymbol{\mu}}_j)(\boldsymbol{z}_j - \hat{\boldsymbol{\mu}}_j)^T, \tag{5.7}$$

$$\text{probabilité } a \text{ priori} \quad : \quad \hat{P}(C_j) = \frac{n_j}{n}. \tag{5.8}$$

[54] Ces points font partie de l'étalon fourni *a priori*, appelé ensemble d'apprentissage, et ne sont pas les mêmes que les points $\boldsymbol{x}$ à classer.

Par conséquent, les paramètres reportés dans l'expression à minimiser (5.5) sont la moyenne empirique $\hat{\boldsymbol{\mu}}_j$ de la classe (le barycentre), la matrice de variance empirique $\hat{\boldsymbol{\Sigma}}_j$ et l'effectif rapporté à la taille n de l'ensemble d'apprentissage ($\hat{P}(C_j)$).

Algorithme de classification non-supervisée

Comme nous l'avons précisé précédemment, l'algorithme que nous proposons consiste en principe en deux étapes (figure 5.18) :

1. **Initialisation.** Un algorithme de classification hiérarchique, assorti d'une technique de découpage du dendrogramme, est utilisé pour séparer l'ensemble des points en classes sous-optimales.
2. **Optimisation.** Une **première phase** itérative consiste en :
   (a) l'élimination des points aberrants de chaque classe ;
   (b) la fusion des classes proches ;
   (c) l'estimation au sens du maximum de vraisemblance des paramètres $\hat{\boldsymbol{\mu}}_j$ et $\hat{\boldsymbol{\Sigma}}_j$ et $\hat{P}(C_j)$ pour les classes obtenues après les pas (a) et (b), suivie de la réallocation des points par analyse discriminante (minimisation selon j de l'expression (5.5)) ;

   Les pas (a)–(c) sont répétés jusqu'à la convergence de la classification.

   Une **deuxième phase** élimine les points aberrants de chaque classe, afin de rendre les classes proches du modèle gaussien choisi (car l'analyse discriminante effectuée à l'étape précédente de l'algorithme place *tous* les points dans la classe la plus probable, même si cette probabilité est très faible).

Initialisation. Il est nécessaire d'initialiser correctement cet algorithme, sans pour autant avoir de connaissances suffisantes pour imposer le nombre, la position ou l'étendue des classes. Pour cela, comme évoqué précédemment, nous suivons l'approche proposée par Fraley et Raftery [45] : les auteurs utilisent les classes fournies par un algorithme hiérarchique pour initialiser un algorithme de partitionnement probabiliste[55].

Un algorithme hiérarchique semble être adapté à notre problème mais, en principe, il faut trouver une méthode correcte d'association des classes afin d'obtenir un dendrogramme représentatif des données (avec le meilleur coefficient cophenetique possible). Les hypothèses de travail précédentes limitent les choix possibles : nous cherchons des classes hyper-ellipsoïdales capables de fournir des estimations initiales fiables des paramètres recherchés (moyennes et variances). La méthode d'association minimale (*single linkage*) peut donc être éliminée, car elle favorise les chaînages des points et la création de classes de formes irrégulières. Il est difficile de choisir entre les autres méthodes hiérarchiques, et nous optons pour la suggestion de Saporta [100] : dans un espace euclidien, la méthode de Ward est adaptée[56].

Il est nécessaire de déterminer un seuil pour le découpage du dendrogramme. Comme nous ne voulons pas imposer de classes de volumes similaires, nous choisissons une méthode de découpage aux différents niveaux basée sur les coefficients d'inconsistance : les branches ayant un coefficient

[55] Cette approche peut être vue inversement comme une méthode pour améliorer les résultats d'un algorithme hiérarchique.

[56] Cependant, vu que la classification hiérarchique fournit seulement une initialisation des classes et que le résultat final est obtenu après une étape probabiliste d'optimisation, les différences entre les méthodes sont peu importantes. En effet, sur l'ensemble réel des points que nous souhaitons classifier, l'algorithme converge vers des solutions très proches indifféremment de l'initialisation hiérarchique choisie.

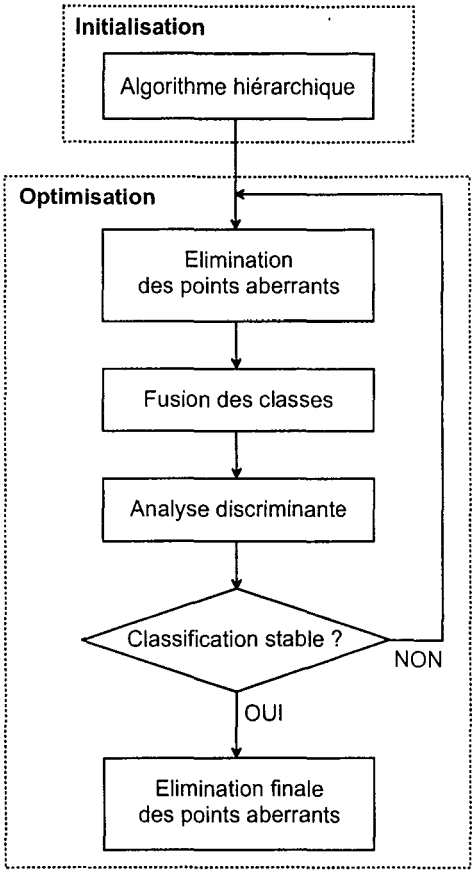


FIG. 5.18 – Organigramme de l'algorithme de classification non-supervisée.

suffisamment grand sont coupées. Reste à définir ce que l'on entend par "suffisamment grand". Notre choix porte sur une signification statistique : un coefficient d'inconsistance suffisamment grand a une valeur significativement différente des autres. Sur une distribution gaussienne (ou même sur une distribution connue quelconque), il est possible de calculer les valeurs significativement éloignées de la moyenne (pour une gaussienne, une valeur située à trois fois l'écart-type de la moyenne peut être vue comme une valeur aberrante, autrement dit différente des autres). Ce n'est pas le cas des coefficients d'inconsistance calculés sur le dendrogramme réel (figure 5.19). Nous avons donc choisi une statistique non-paramétrique, similaire à la médiane et appliquée sur les coefficients différents de 0 : le seuil est égal à la valeur de percentile=99 (99% des valeurs lui sont inférieures).

Optimisation. La classification obtenue à l'étape précédente n'est certainement pas optimale : les classes obtenues par la méthode de Ward ont tendance à être sphériques et de variances proches, et leur nombre peut être très important, en fonction du seuillage des coefficients d'inconsistance choisi (voir figure 5.20). L'objectif premier de l'optimisation est donc l'amélioration de la classification hiérarchique, dans le sens du modèle que nous avons choisi : classes ellipsoïdales de volumes et densités quelconques.

Il est à notre avis important de rappeler ici l'objectif final de notre étude : l'analyse de l'activité gastro-intestinale par la classification des séquences sonores et des sons abdominaux individuels. Cette classification doit être supervisée, dans un but diagnostique. Autrement dit,

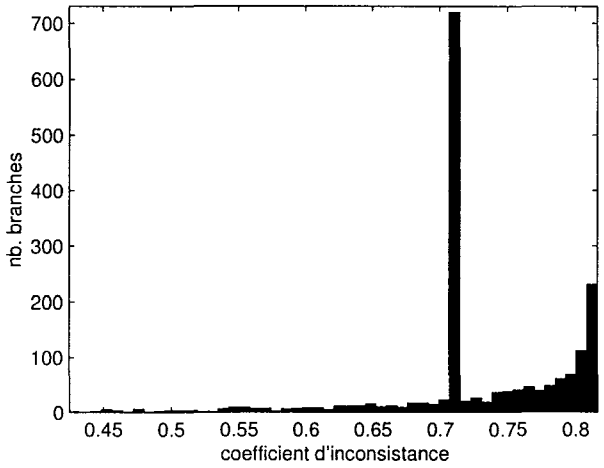


FIG. 5.19 – Exemple d'histogramme de coefficients d'inconsistance, calculés sur un dendrogramme obtenu par la méthode de Ward sur l'ensemble des séquences de phonoentérogramme. L'axe des abscisses est limité inférieurement, donc le pic en 0 dû aux branches situées sur le premier niveau du dendrogramme n'est pas présenté. Le pic en 0,71 est dû aux branches immédiatement supérieures, qui relient une classe formée de deux points à un troisième point.

l'objectif de l'algorithme de classification non-supervisée est de faciliter une étape *supervisée* applicable pour l'étude de la physiologie abdominale par la création d'un ensemble d'apprentissage. Les caractéristiques des classes constituant cet ensemble dirigent le choix de l'algorithme de classification supervisée : en effet, une façon classique et qui fournit de bons résultats est l'analyse discriminante que nous avons présentée précédemment (page 139). Par conséquent, nous pouvons fixer un deuxième objectif pour l'étape d'optimisation, cette fois méthodologique : elle doit fournir un ensemble d'apprentissage adapté et compatible avec l'analyse discriminante. Pour être précis, notons que l'analyse discriminante a besoin d'estimations fiables des paramètres des lois de densité (équations 5.6 à 5.8), ce qui impose des conditions sur la construction des classes qui constitueront l'ensemble d'apprentissage. De plus, il est naturel de demander à cet ensemble d'être stable par rapport à la méthode supervisée qui l'utilise : si les mêmes séquences ou sons sont présentés à l'algorithme de classification, ils doivent être placés dans les mêmes classes.

Détaillons les inconvénients de la solution hiérarchique qui doivent être abordés lors de l'optimisation :

- l'histogramme présenté figure 5.20 montre l'effectif des classes obtenues par classification hiérarchique sur l'ensemble des séquences de phonoentérogramme que nous avons traitées. Il est évident que, du point de vue médical, l'information est trop éparpillée : il n'y a pas 100 types d'activité sonore abdominale distinctes du point de vue pratique ;
- un autre inconvénient important, lié au grand nombre de classes, est l'effectif très faible de certaines d'entre elles : il est inutile de créer une classe spécifique pour deux séquences de longueur d'une minute sur plus de 3000 minutes d'enregistrement. Il y a une autre dimension de ce problème, liée à l'algorithme de classification supervisée : afin d'obtenir une matrice de variance d'une classe non-singulière, on a besoin d'au moins $d+1$ points, où d est la dimension de l'espace [65]. Qui plus est, pour obtenir une estimation correcte des paramètres des densités de probabilité (équations 5.6 à 5.8), il est nécessaire d'avoir un nombre minimal de points. Ce nombre est évalué à $5 \cdot d+1$ ou à $10 \cdot d+1$ dans [61, 79] (16, respectivement 31 points dans un espace tridimensionnel), ou encore à environ $10^{0,6 \cdot (d-0,25)}$ dans [60] (≈ 44 points en 3D) ;

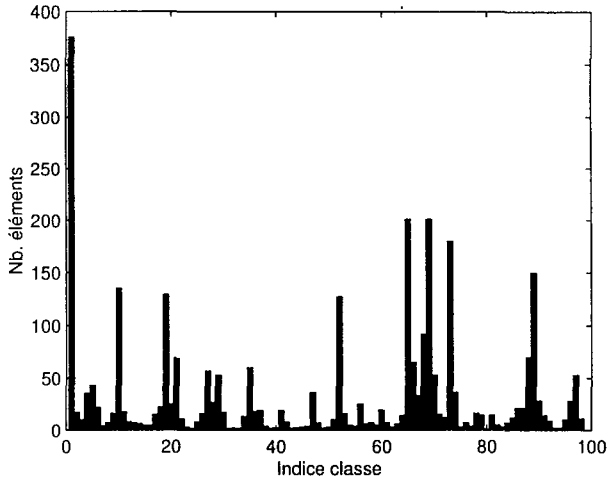


FIG. 5.20 – Histogramme des séquences de phonoentérogramme par classes, après l'étape de classification hiérarchique par méthode de Ward et découpage du dendrogramme par seuillage des coefficients d'inconsistance. On obtient 98 classes d'effectifs variant entre 2 et 378 points.

- la classification hiérarchique ne prend pas en considération la notion de point aberrant, tous les points de l'espace étant attachés à une classe plus ou moins nombreuse. Il est peut-être intéressant de détecter ces valeurs aberrantes (en fait des séquences ou des sons individuels rares), et de les soumettre séparément au jugement médical. De plus, du point de vue méthodologique, les points aberrants perturbent les estimations des paramètres des classes ;
- il existe des classes proches l'une de l'autre, voire en partie superposées. Il est probable que ces classes n'aient pas de signification médicale distincte et qu'elles puissent être fusionnées.

Afin de résoudre les problèmes listés ci-dessus, l'optimisation se déroule en plusieurs étapes :

(a) **la détection des points aberrants**, classe par classe. Cette étape nécessite de préciser ce que l'on entend par points aberrants : ils sont définis comme les éléments dont la probabilité d'appartenance à la classe est trop faible. Ceci n'est pas encore une définition précise, car il faut spécifier quel est le seuil inférieur acceptable pour la probabilité d'appartenance.

Pour une densité gaussienne, l'approche classique [93] est d'estimer la probabilité d'appartenance à partir de la distance de Mahalanobis (5.1) entre le point et le barycentre : dans un espace tridimensionnel, cette distance suit une loi de χ^2 à 3 degrés de liberté [100]. En définissant les points aberrants comme ceux dont la probabilité est inférieure à $\alpha = 0,01$ ou $\alpha = 0,05$, on peut les détecter par un test de χ^2 et calculer ainsi un *seuil d'exclusion*. Comme la distance de Mahalanobis est fondée sur l'estimation de la matrice de variance de la classe, nous appliquons cette technique seulement pour les classes ayant un effectif supérieur à $10 \cdot d + 1$, où d est la dimension de l'espace. Tous les points appartenant à des classes d'effectif inférieur seront aussi considérés comme aberrants.

Selon Rousseeuw [93], si cette méthode est fiable pour la détection d'une seule valeur aberrante, elle est perturbée par la présence d'un nombre plus important. Une solution possible est de détecter ces valeurs itérativement, jusqu'à leur identification complète[57]. C'est la méthode que nous avons adoptée, le résultat obtenu étant la création de "noyaux" robustes de classes ;

[57]Remarquons à l'occasion que cette méthode est basée sur le même principe que l'algorithme de débruitage décrit dans le troisième chapitre de cette thèse : la détection itérative des valeurs "significatives".

La classification présentée figure 5.21(a) n'a pratiquement pas d'utilité : une grosse partie des séquences sont considérées aberrantes (plus de 34,1%), la majorité sont regroupées dans une seule classe (no. 1, 59,2%) et la seule classe potentiellement intéressante dans cette configuration (no. 2, 6,7%) contient en fait les séquences complètement silencieuses (autrement dit, c'est une classe infiniment dense, car elle occupe un seul point de l'espace, et elle est facilement retrouvée par tout algorithme de classification ou encore par simple inspection du signal).

La deuxième classification (figure 5.21(b)) semble plus utile. Un seuil de fusion nul empêche toute fusion des classes et réduit le nombre de points aberrants (9,1% de l'ensemble). La classe silencieuse est bien identifiée (classe no.4). Cette classification révèle des informations intéressantes quand on l'étudie en détail, patient par patient (figure 5.22) ou voie par voie (figure 5.23).

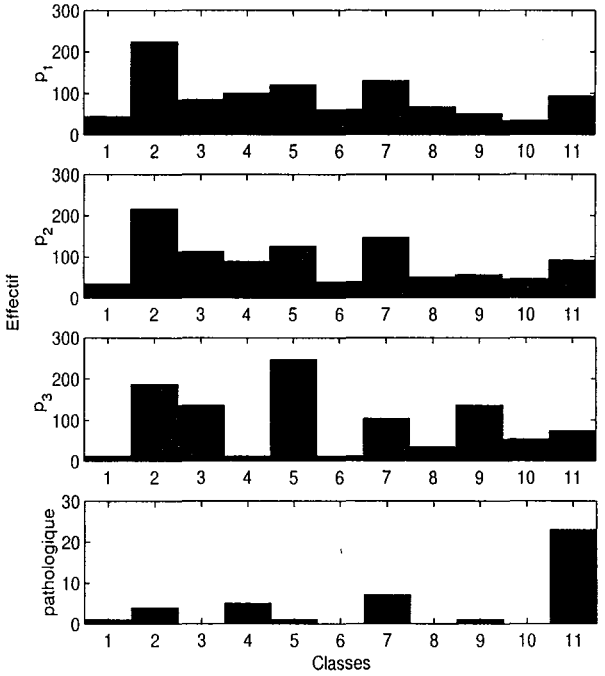


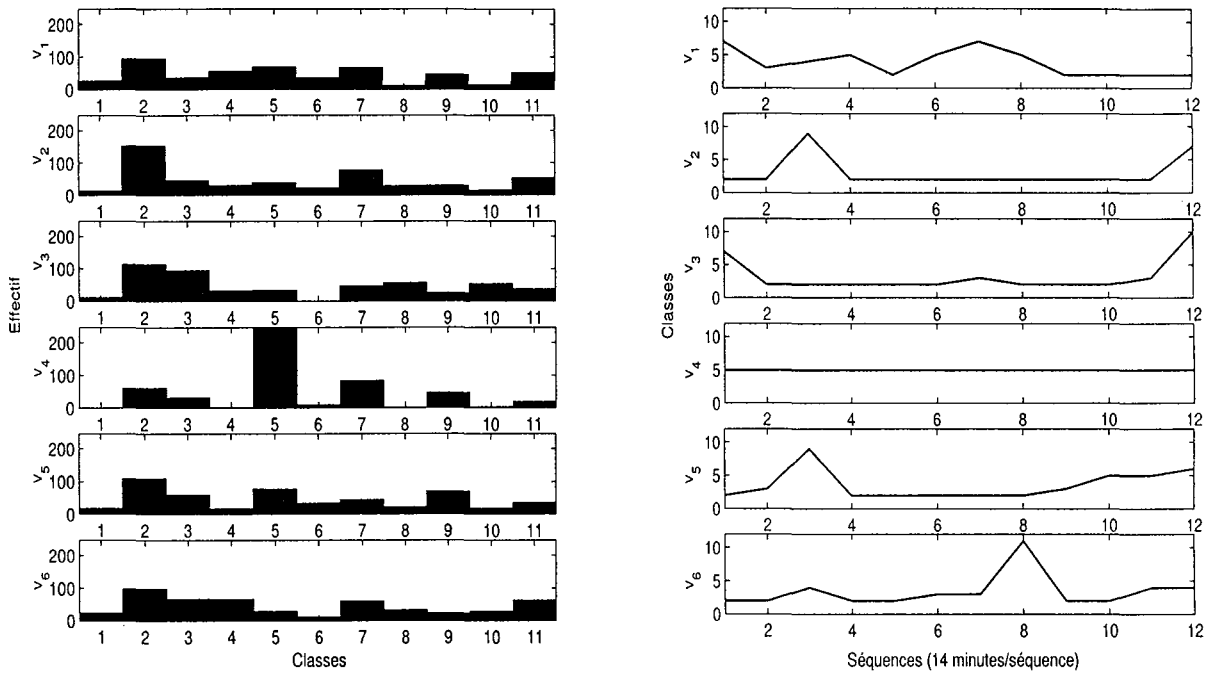
FIG. 5.22 – Répartition par classes des séquences des trois patients sains et du patient malade (l'échelle du dernier histogramme à été modifiée, vu que la durée du phonoentérogramme pathologique est plus courte et donc le nombre d'événements moins important).

Pour les phonoentérogrammes sains p_1, p_2 et p_3, on obtient respectivement 9,1%, 9,0% et 7,3% de séquences "aberrantes". En revanche, pour le signal pathologique, ce pourcentage devient 54,8%. On note également que la classe no. 5 est très bien représentée chez le patient p_3, mais que la distribution des séquences dans les autres classes est relativement similaire pour les trois patients sains (avec un bémol pour la classe no. 9, également plus représentée chez p_3). Si on essaie de trouver la classe la plus caractéristique d'un patient (mis à part le cas pathologique, représenté par un fort pourcentage de points aberrants), alors le patient p_3 est représenté par la classe no. 5 (24,5%), alors que les patients p_1 et p_2 sont représentés par la classe no. 2 (24,5%, respectivement 21,5% des points). Il est clair que, vu les pourcentages, cette représentativité est faible et que l'on ne peut pas réduire un patient à une seule classe (sauf cas pathologique).

Une information supplémentaire est donnée par les histogrammes des classes par voie (figure 5.23), où l'on remarque une très forte présence de la classe no.5 sur la voie v_4 (49,0%), alors que les autres voies ont une activité prépondérante de type 2 (classe no.2). Cette représentation masque l'évolution temporelle des signaux, car on peut imaginer une "trajectoire" des

phonoentérogrammes selon les classes trouvées (analogue à l'évolution selon une des composantes principales présentée figure 5.10, page 128) : pour chaque voie, on considère la classe la plus représentative séquence après séquence et on trace la courbe obtenue. Cette trajectoire est présentée figure 5.23(b) : elle montre sans aucune ambiguïté que la classe 5 est vraiment représentative de la voie v_4, mais également que la représentativité de la classe no. 2 pour les autres voies dépend du temps. Par exemple, sur la voie v_1, une seule séquence parmi les neuf premières est de type 2.

Quelle interprétation donner à ces constats ? Est-il possible d'envisager une différenciation entre l'activité des différents organes de la cavité abdominale ? La voie v_4 est la plus représentative de l'intestin grêle, grâce à son positionnement sous-ombilical (voir figure 2.4, page 26), les voies v_3, v_2, v_6 et v_5 suivent le trajet du côlon et la voie v_1 est placée à la fois sur l'estomac et le côlon. Une hypothèse qui reste à vérifier serait donc la suivante : l'activité caractéristique de l'intestin grêle est de type 5 et celle du côlon plutôt de type 2, mais elle change après environ 2 heures 30 après le repas (ou juste avant le repas suivant, quand la sensation de faim commence à être ressentie). L'estomac a une activité variable pendant la période postprandiale, celle-ci masque l'activité spécifique du côlon, et il est plutôt silencieux une fois la digestion gastrique terminée.



(a) histogrammes

(b) évolution

FIG. 5.23 – Répartition (a) et évolution (b) en classes des séquences, voie par voie.

Les résultats obtenus en appliquant l'algorithme de classification décrit aux sons individuels sont difficile à interpréter et ils varient beaucoup avec les paramètres. Pour un seuil de fusion haut, on arrive même à regrouper les sons dans une seule classe et un nuage de points aberrants, alors que pour un seuil nul, on obtient environ 30 classes, dont une partie très semblable à l'écoute. Il est donc nécessaire de tester davantage la méthode, et sur une base de données plus importante, avant d'essayer une analyse pertinente. Ces travaux sont relégués en perspective.

5.3 Conclusion, perspectives

Ce chapitre attaque le dernier volet du projet d'étude des sons abdominaux : l'analyse des signaux. Nous avons envisagé deux types d'analyse : globale et locale.

L'analyse globale considère le signal de phonoentérographie comme une succession de séquences de durée fixe (1 minute), caractérisées par différents indices d'activité. Les séquences peuvent être représentées comme des vecteurs ou des points dans un espace engendré par les indices d'activité.

Ces indices, dont le but est de fournir une description pratique et interprétable médicalement, sont calculés à partir des grandeurs physiques mesurables comme l'énergie, la fréquence, le nombre d'événements, etc. Les sources bibliographiques consultées utilisent plusieurs indices différents, dont une partie sont vraisemblablement redondants. Notre première approche a visé l'étude de ces indices afin d'éliminer la redondance et d'optimiser la représentation des séquences de phonoentérogramme en réduisant le nombre des caractéristiques nécessaires pour les décrire. Nous avons utilisé pour cela l'analyse en composantes principales, qui nous a permis de caractériser les phonoentérogrammes à l'aide de trois ou quatre nouvelles variables indépendantes. L'espace ainsi crée préserve l'essentiel de la dispersion des points (séquences), tout en gardant une signification physique des nouvelles variables explicatives (composantes principales).

Dans l'espace des composantes principales, nous avons ensuite analysé les différents groupes de séquences. Ainsi, nous avons pu conclure (avec toute la réserve imposée par une base de données insuffisamment développée), que cet espace et ces nouvelles variables sont adaptés à la description des phonoentérogrammes et que cette méthode est capable de détecter les différences de fonctionnement entre les endroits de l'abdomen, au cours du cycle digestif et entre les patients sains et malades.

Sachant que l'analyse en composantes principales n'est pas particulièrement conçue pour discriminer les différents modes de fonctionnement, nous avons essayé d'améliorer les résultats obtenus par une méthode de classification. Cependant, nous nous sommes heurtés au problème essentiel soulevé par l'absence d'un étalon de classification qui puisse servir comme une base d'apprentissage. Nous avons essayé de créer cet étalon en proposant une méthode de classification non-supervisée, censée trouver des regroupements naturels de séquences de phonoentérogramme. A la différence de l'analyse en composantes principales, qui est essentiellement une méthode paramétrique car elle différencie un nombre connu de groupes représentés par leur moyennes, la classification non-supervisée cherche à estimer à la fois le nombre de classes et leurs paramètres. Les résultats que nous avons obtenus sont difficilement interprétables et validables en l'absence d'un étalon et il n'est pas certain que la méthode que nous avons développée fonctionne dans le cadre qui nous intéresse. Cette validation, sur une base de signaux plus importante, fait partie de nos perspectives de travail.

Nous avons suivi une démarche similaire pour **l'analyse locale** des phonoentérogrammes. Dans ce cadre, un signal n'est plus interprété comme une suite de séquences, mais comme une collection de sons abdominaux isolés. Ces sons sont caractérisés également par des grandeurs physiques énergétiques, fréquentielles et temporelles. Nous avons obtenu, par analyse en composantes principales, un espace de représentation de dimension réduite et avec des variables physiquement interprétables. Vu que les mêmes grandeurs physiques sont à la base de la construction de l'espace de représentation des séquences et des sons individuels, les deux représentations sont en partie redondantes. Comme l'optique générale de cette thèse est la construction d'une boîte à outils adaptés au traitement et à l'analyse des phonoentérogrammes, et que la base de signaux nécessaires pour une validation correcte de ces outils (et pour leur amélioration) n'est pas encore disponible, nous ne proposons pour l'instant aucun choix.

Conclusion générale. Perspectives

Les sons physiologiques ont un potentiel diagnostique important et leur auscultation fait partie intégrante de l'examen clinique. Leur contenu informatif est reconnu et exploité surtout dans le cadre de l'auscultation cardiaque et pulmonaire. En revanche, l'auscultation abdominale est moins développée, car les difficultés liées à l'acquisition et à la nature des signaux (information répartie sur des longues durées, apparition irrégulière et caractéristiques non-stationnaires des sons) rendent l'interprétation subjective et ambiguë. Pour pallier ces difficultés, plusieurs travaux proposent des méthodes d'enregistrement des sons abdominaux, suivies de méthodes d'interprétation plus ou moins automatiques. Les résultats présentés dans la littérature semblent confirmer l'utilité générale de cette méthode d'investigation dans la surveillance post-opératoire, dans la recherche pharmacologique, voire même dans le diagnostic de certaines maladies fonctionnelles. Cependant, les approches ne sont pas systématiques et les résultats sont parfois contradictoires.

La question que nous nous sommes posée au début de nos recherches a été simple : peut-on concevoir un outil d'auscultation abdominale automatisée fiable et utilisable systématiquement dans la recherche physiologique, voire dans l'examen clinique ? Trouver une réponse à cette question implique une recherche dont l'objectif est assez ambitieux : développer des méthodes pour trouver, extraire et analyser l'information contenue dans les signaux sonores abdominaux. Cet objectif se traduit par la construction d'une "boîte à outils" contenant les méthodes et les algorithmes nécessaires. Cette recherche méthodologique de traitement du signal est complétée par une recherche technologique, menée sur la partie instrumentation.

La première étape pour atteindre l'objectif énoncé a été une étude bibliographique sur l'objet de nos recherches, à savoir les manifestations sonores normales et pathologiques de l'appareil digestif. Cette étude est exposée dans le premier chapitre du mémoire, consacré à la présentation du contexte médical, mais aussi aux méthodes d'acquisition, de traitement et d'interprétation des phonoentérogrammes. Cet état de l'art a permis de définir la problématique scientifique de nos travaux, qui nous a amené à nous intéresser à des domaines relativement différents comme l'instrumentation, le pré-traitement (détection, débruitage, segmentation et localisation) et l'analyse de données (analyse en composantes principales et classification).

Ainsi, nous avons conçu une chaîne d'instrumentation adaptée à l'acquisition des phonoentérogrammes, décrite dans le deuxième chapitre de ce mémoire. Ses caractéristiques offrent un certain nombre d'avantages :

- le capteur, constitué d'une tête de stéthoscope et d'un microphone, assure une atténuation partielle des bruits ambiants parasites et permet une validation médicale directe, car les sons fournis sont proches de ceux rencontrés en auscultation clinique ;
- le système multi-voies permet le placement de capteurs en plusieurs endroits de l'abdomen et donc une acquisition en parallèle qui facilite la localisation des sons ;
- l'enregistreur numérique permet la numérisation et le stockage de grandes quantités de données, autrement dit de signaux de longue durée, indispensables pour une analyse de l'activité abdominale.

En collaboration avec l'INRS, nous avons étudié les caractéristiques de l'instrumentation proposée et vérifié sa fiabilité par rapport aux changements dans les conditions d'utilisation, notamment la variation de la pression d'appui de la tête de stéthoscope sur la peau.

Le pré-traitement que nous proposons dans le troisième chapitre cherche d'abord à prendre en compte la nature non-stationnaire des signaux. Nous présentons les différentes méthodes de détection, débruitage et segmentation des signaux non-stationnaires, en mettant un fort accent sur les méthodes temps-échelle. L'algorithme itératif de détection et débruitage par ondelettes par une méthode de point fixe, que nous introduisons et développons dans ce chapitre, permet une optimisation importante au niveau de la charge de calcul, par rapport aux autres algorithmes itératifs présentés dans la littérature. A partir d'une modélisation probabiliste de la distribution des coefficients d'ondelette par des gaussiennes généralisées, nous proposons plusieurs versions de paramétrisation automatique de l'algorithme de débruitage. Nous comparons les performances des différentes versions de notre algorithme aux algorithmes classiques du domaine, sur une batterie de signaux simulés et réels. Selon ces tests, notre algorithme se révèle plus performant que les autres sur certains types de signaux et il est notamment adapté au pré-traitement des phonoentérogrammes. Une partie de ces résultats fait l'objet de plusieurs travaux publiés dans une conférence internationale [87] ou en revues [86, 89]. L'algorithme est également en phase de test pour une toute autre application : la détection des bruits parasites dans les habitacles de voiture (travaux réalisés par la société 01dB Acoustics & Vibration - Metravib Technologies [14]).

Le quatrième chapitre enrichit le pré-traitement antérieur par une approche pragmatique qui introduit des connaissances *a priori* sur les caractéristiques physiques des sons abdominaux. Ces caractéristiques physiques (durée, fréquence principale, intensité acoustique) sont obtenues à partir du vecteur de coefficients d'ondelette de chaque événement, dont nous présentons une méthode de construction à partir du vecteur de coefficients par la décomposition du signal entier. Les caractéristiques de chaque événement détecté et segmenté par l'algorithme de débruitage par ondelettes sont comparées aux connaissances *a priori* sur les sons, ce qui permet l'élimination des artéfacts et la re-segmentation des signaux.

La deuxième partie du quatrième chapitre introduit l'information apportée par les enregistrements multi-voies. Nous proposons et justifions, dans cette partie, une méthode empirique de localisation de sources sonores, après avoir testé les méthodes de localisation décrites dans la littérature et montré leur inadaptation à notre cas. Nous pensons que la localisation de sources sonores à l'intérieur de l'abdomen est difficilement réalisable, car le milieu de propagation est très complexe et ses caractéristiques varient dans le temps. Cette méthode aboutit à une nouvelle re-segmentation des phonoentérogrammes, car elle détecte et élimine les sons propagés. Les travaux sur la localisation ont également été publiés dans [90].

Ce chapitre peut également être vu comme une préparation des phonoentérogrammes pour l'étape finale d'analyse, car il introduit les caractéristiques physiques qui servent à décrire chaque événement[58] ainsi que les indices d'activité adaptés à la description de séquences de phonoentérogramme.

Le cinquième et dernier chapitre s'attaque au vaste domaine de l'analyse de données. Dans le cadre de notre application, les données sont soit des séquences de phonoentérogramme soit

[58]En fait, dans une optique de boîte à outils, nous proposons une méthode pour obtenir les coefficients d'ondelette de chaque événement. Ces coefficients seront utilisés, dans la perspective de nos travaux, pour une classification des sons plus précise que celle que nous présentons ici.

des sons individuels, décrits respectivement par des indices d'activité et par des caractéristiques physiques. Nous proposons une méthode de construction d'espaces adaptés à la représentation de ces données basée sur l'analyse en composantes principales. Nous interprétons ensuite dans ces espaces les nouvelle variables (les composantes principales) et les données (séquences de phonoentérogramme / sons individuels). La méthode fournit des résultats intéressants et elle arrive à différencier les types d'activité spécifiques à plusieurs endroits de l'abdomen, les phases du cycle digestif ou encore des patients sains et malades.

Afin d'améliorer les débuts de classification offerts par l'analyse en composantes principales, nous proposons dans la dernière partie de cette thèse un algorithme de classification non-supervisée censé trouver des classes naturelles dans l'ensemble de séquences de phonoentérogramme, respectivement de sons individuels. Les résultats sont difficiles à interpréter (car la base de données est réduite), mais ils sont intéressants et prometteurs et ils confirment et améliorent les résultats obtenus après analyse en composantes principales. On arrive notamment à faire la distinction entre les patients sains et malades, mais aussi à associer un type d'activité à une localisation, voire à un organe particulier. Les premiers essais de classification de sons abdominaux ont fait l'objet d'une communication en congrès [88].

Comme nous l'avons dit ci-dessus, l'objectif fixé est ambitieux. Les méthodes que nous présentons dans ce mémoire, même si elles s'attaquent à tous les volets du problème du traitement des phonoentérogrammes, ne sont pas encore prêtes pour une expérimentation clinique. Néanmoins, les résultats sont encourageants et ouvrent des perspectives de recherche intéressantes.

Le protocole d'acquisition que nous avons utilisé montre l'intérêt des enregistrements de longue durée et sur plusieurs voies. Nous avons détecté des différences d'activité au cours des 168 minutes d'enregistrement après un repas standardisé. Cependant, les différences inter-individuelles sont assez importantes, ce qui suggère (outre le fait que chaque individu a une physiologie spécifique) que les repas précédents influencent aussi le fonctionnement de l'appareil digestif (le transit alimentaire peut durer plusieurs jours). Dans une optique de recherche physiologique (et non pas en application clinique), il serait intéressant de standardiser la prise alimentaire sur des durées plus longues et d'adapter le protocole d'enregistrement en le découpant en plusieurs tranches de durée limitée mais réparties dans le temps.

Il serait souhaitable également de définir plus précisément les repères anatomiques qui déterminent les positions des capteurs, vu l'importance du positionnement pour la localisation et donc pour toute l'analyse de l'activité abdominale qui suit.

Ces modifications du protocole peuvent conduire à repenser l'instrumentation : le positionnement plus strict et les durées d'enregistrement plus longues seraient facilités par des capteurs plus légers, dont la fixation sur la peau de l'abdomen est plus facile.

Une amélioration nécessaire du pré-traitement est l'introduction du filtrage adaptatif, dont l'objectif sera à la fois l'atténuation du bruit ambiant et celle des autres bruits physiologiques (cardiaques et pulmonaires).

L'algorithme de débruitage que nous avons présenté est assez général et nous avons montré qu'il est adapté au traitement des phonoentérogrammes. Ses performances méritent d'être testées et comparées aux performances fournies par d'autres algorithmes sur d'autres types de signaux réels, afin de juger de son utilité et de pouvoir le recommander pour des applications spécifiques.

Des travaux de recherche importants restent à faire surtout dans le domaine de l'analyse et de l'interprétation des phonoentérogrammes. L'amélioration la plus importante que nous imaginons pour l'analyse globale implique le découpage des signaux en séquences (fenêtres d'observation

et de calcul des indices d'activité) de durée variable. La durée de ces fenêtres peut et doit être modifiée afin de rechercher d'autres informations, disponibles sur plusieurs niveaux de résolution :

- le rythme d'apparition des contractions gastriques et/ou intestinales est inférieur à une minute, donc la fenêtre d'observation doit tenir compte de ce fait – elle doit être plus courte et peut-être glissante dans le temps ou en tout cas, avec un pas de déplacement inférieur à sa taille ;
- la durée des phases de digestion (les phases du complexe migrant moteur, par exemple) est supérieure à la minute, et l'utilisation de fenêtres d'observation plus longues et/ou glissantes (et donc le lissage des indices d'activité) peut également faciliter leur détection.

En ce qui concerne l'analyse des phonoentérogrammes à l'aide des sons individuels, il est clair que la méthode que nous proposons est assez rudimentaire et que, de plus, l'information qu'elle fournit est en partie redondante avec l'analyse globale. L'explication vient principalement des caractéristiques choisies pour la description des sons : elles sont les mêmes que celles utilisées pour le calcul des indices d'activité et elles ne gardent qu'une information basique sur les sons (durée, intensité, fréquence). Cependant, la méthode que nous avons présentée dans le quatrième chapitre permet l'extraction des coefficients d'ondelette de chaque événement et donc une caractérisation beaucoup plus fine. On peut donc envisager l'utilisation directe de ces coefficients (ou des grandeurs dérivées comme l'énergie par bande fréquentielle/intervalle temporel) pour l'analyse et la classification des sons abdominaux.

Comme nous l'avons précisé dans le cinquième chapitre, le domaine de la classification est très vaste et le nombre de méthodes développées très important. Il serait intéressant de tester plusieurs de ces méthodes sur nos ensembles de données, afin de choisir la plus adaptée. Cependant, il n'est pas certain que la classification non-supervisée puisse retrouver des classes naturelles dans une telle base de données. Il n'est pas sûr, par exemple, qu'une méthode automatique puisse tracer une frontière nette entre les cas normaux et pathologiques. A notre avis, il est plutôt intéressant de construire une base de données déjà classées par un expert, autrement dit un ensemble d'apprentissage qui sera ensuite utilisé pour une classification supervisée (comme par exemple l'analyse discriminante).

Notre expérience nous laisse penser que le traitement et l'analyse des phonoentérogrammes peut être un outil d'investigation de la physiologie abdominale, voire de l'aide au diagnostic. Cependant, il est indispensable que tous ces travaux soient validés par une interprétation médicale qui puisse juger de la pertinence de l'étape finale d'analyse. Enfin, seule l'application dans des situation réelles permettra d'apprécier l'utilité médicale de ces résultats.

Annexes

A

Ondelettes testées

Nous présentons dans cette annexe les ondelettes Daubechies no.9 et Symlets no.9, qui font partie des ondelettes testées dans l'algorithme de débruitage (voir section 3.7.4). L'ondelette choisie est Daubechies no.9.

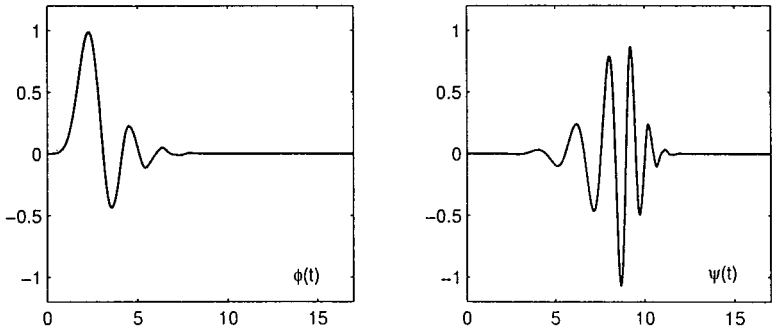


FIG. A.1 – Fonction d'échelle et ondelette Daubechies no.9.

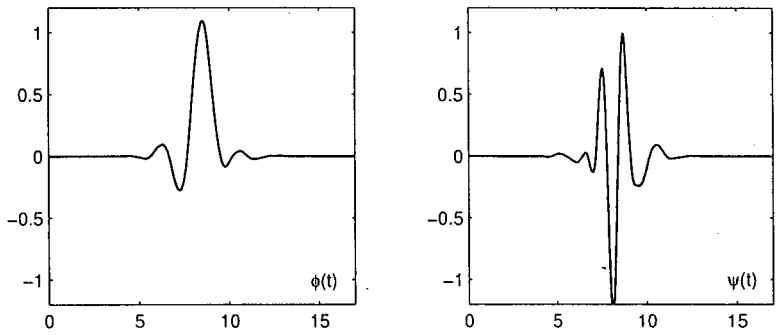


FIG. A.2 – Fonction d'échelle et ondelette Symlets no.9.



Validation algorithmes de débruitage. Résultats détaillés

Nous présentons dans cette annexe les résultats détaillés des étapes de validation décrites dans la section 3.7. Nous avons choisi une présentation graphique, pour faciliter la lecture des résultats et la comparaison des algorithmes de débruitage.

Outre les remarques que nous avons faites dans la section dédiée à la validation des algorithmes, nous pouvons ajouter quelques commentaires :

- concernant l'erreur quadratique moyenne EQM (3.32), les résultats obtenus confirment l'observation sur l'élimination du bruit laplacien (paragraphe 3.7.5, page 72). En effet, pour un même type de signal, les performances semblent légèrement inférieures pour un bruit superposé de distribution laplacienne ;
- la tendance est inversée pour l'erreur de segmentation ES (3.33) ;
- enfin, l'appréciation du nombre d'événements (critère ED (3.34)) dépend aussi de la distribution du bruit – si leur nombre est surévalué et donc les fausses alarmes sont prédominantes pour le bruit laplacien (ce qui se corrèle aux valeurs du critère ES), on peut avoir même des non-détections pour les autres types de bruit.

Enfin, les valeurs des critères dépendent de la puissance et du contenu fréquentiel du bruit. Les performances des algorithmes de débruitage sont plus faibles quand le bruit est fort (ce résultat est bien sûr attendu).

Dans le cas d'un bruit coloré de puissance plus importante dans les basses fréquences, les performances du débruitage sont plus faibles, car les événements (plutôt basse-fréquence aussi) sont relativement noyés.

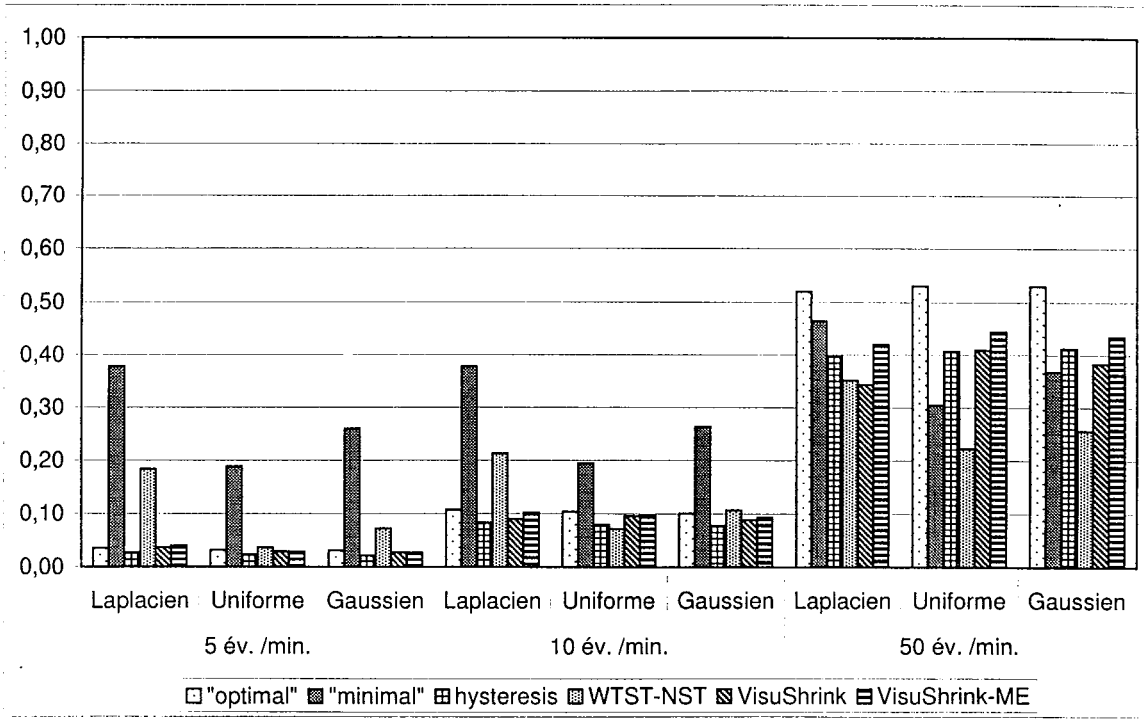


FIG. B.1 – Valeurs détaillées du critère EQM pour les 6 algorithmes testés, pour un bruit additif blanc, en fonction du type de signal et de la distribution du bruit.

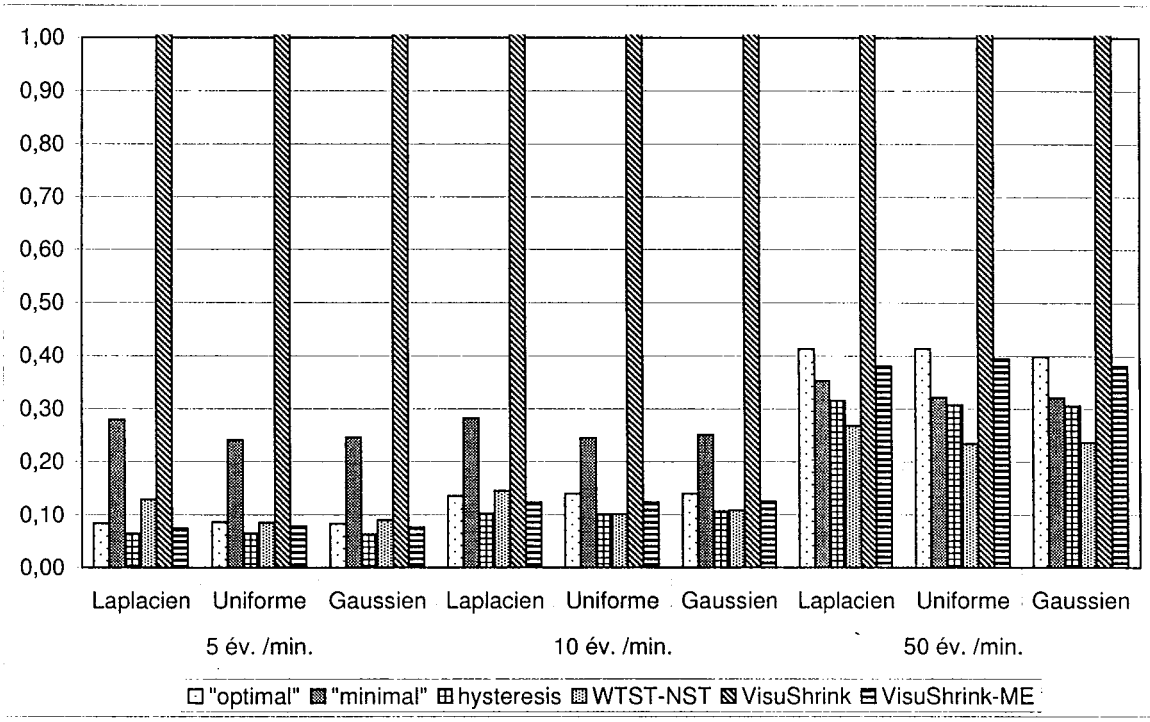


FIG. B.2 – Valeurs détaillées du critère EQM pour les 6 algorithmes testés, pour un bruit additif coloré, en fonction du type de signal et de la distribution du bruit.

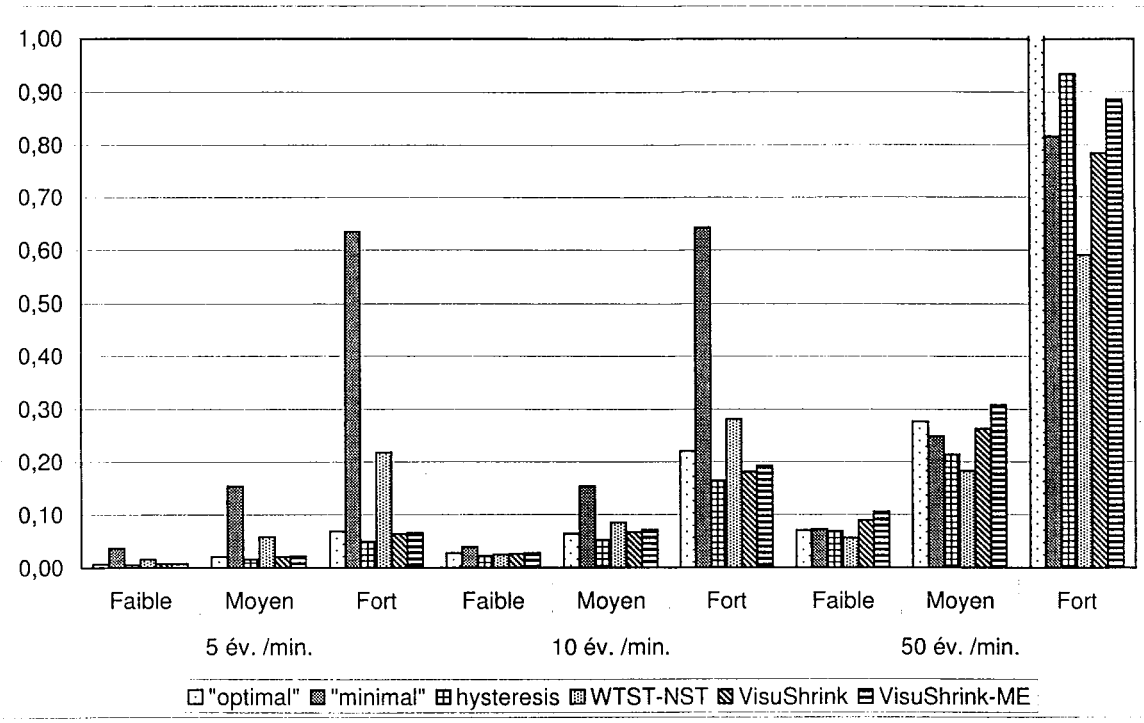


FIG. B.3 – Valeurs détaillées du critère EQM pour les 6 algorithmes testés, pour un bruit additif blanc, en fonction du type de signal et de la puissance du bruit (Faible : $\sigma = 0.5$, Moyen : $\sigma = 1$, Fort : $\sigma = 2$).

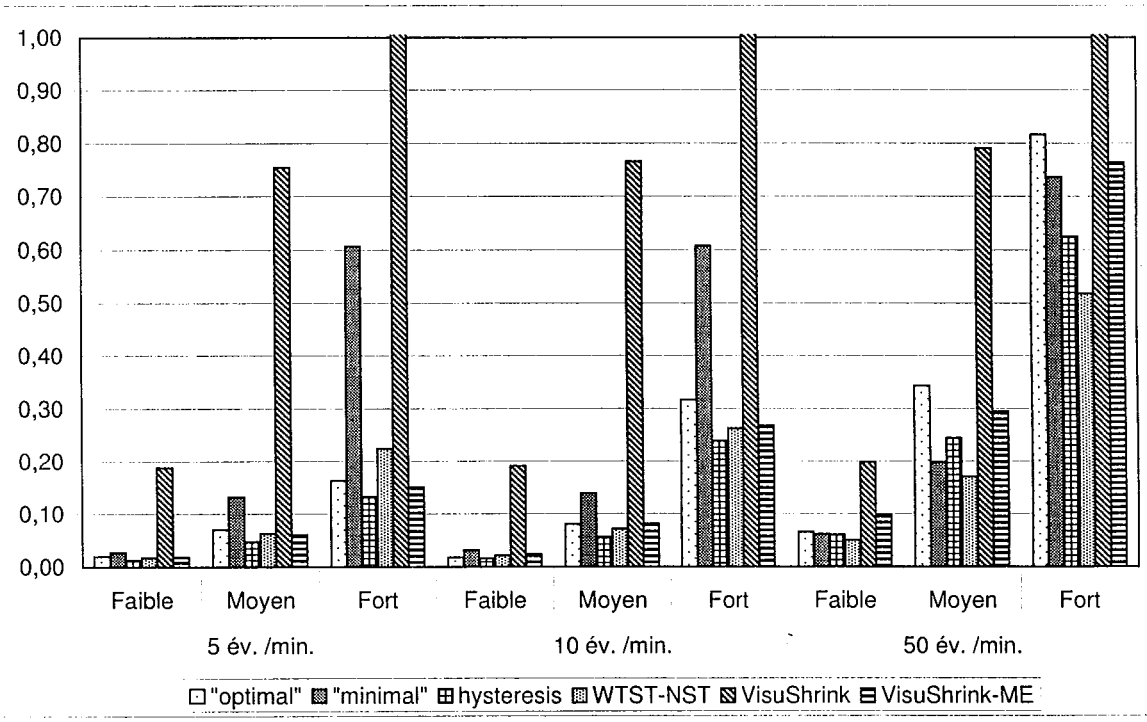


FIG. B.4 – Valeurs détaillées du critère EQM pour les 6 algorithmes testés, pour un bruit additif coloré, en fonction du type de signal et de la puissance du bruit (Faible : $\sigma = 0.5$, Moyen : $\sigma = 1$, Fort : $\sigma = 2$).

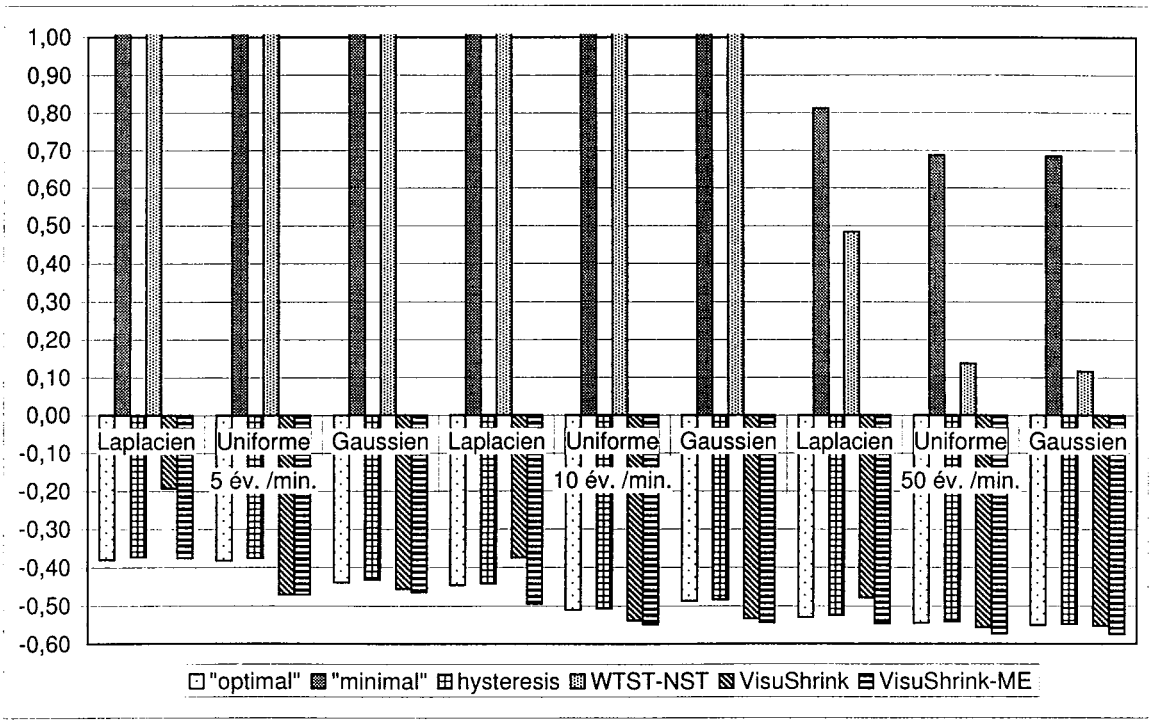


FIG. B.5 – Valeurs détaillées du critère *ES* pour les 6 algorithmes testés, pour un bruit additif blanc, en fonction du type de signal et de la distribution du bruit.

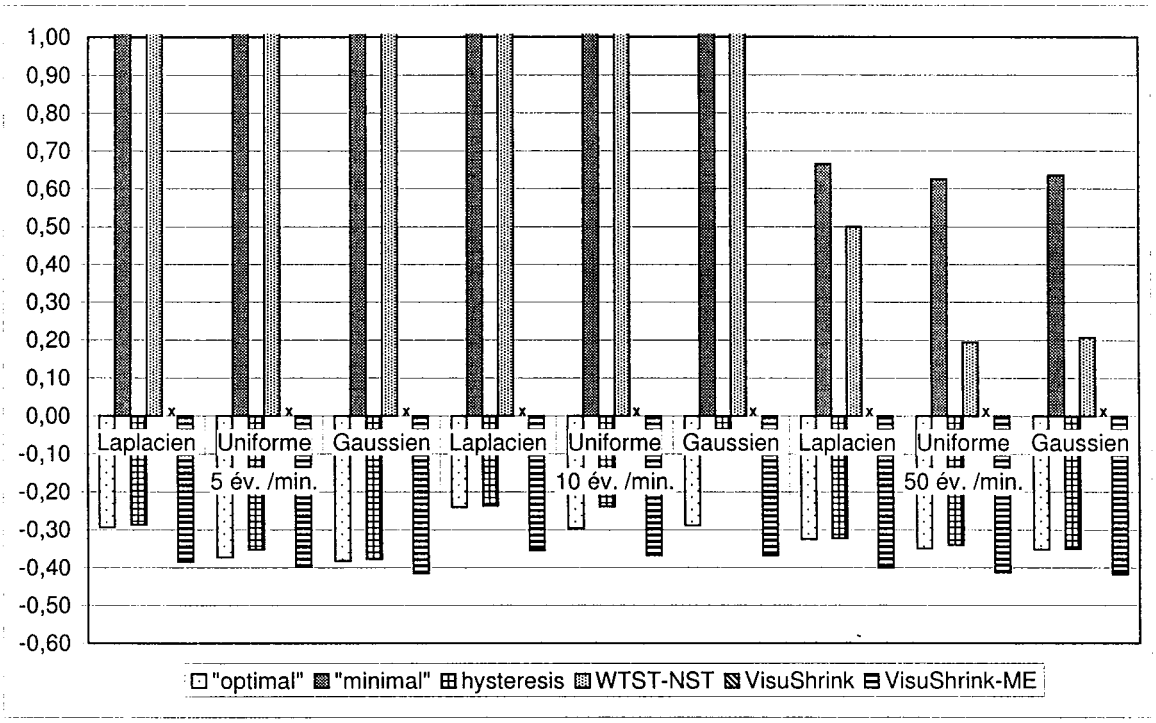


FIG. B.6 – Valeurs détaillées du critère *ES* pour les 6 algorithmes testés, pour un bruit additif coloré, en fonction du type de signal et de la distribution du bruit.

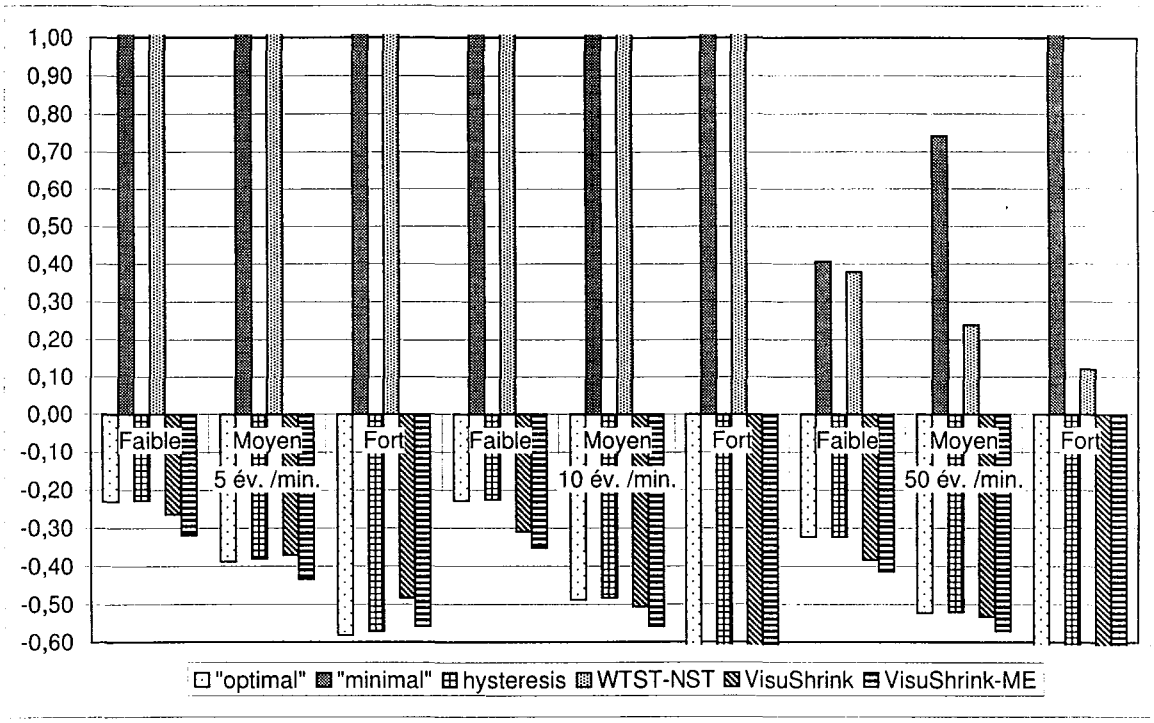


FIG. B.7 – Valeurs détaillées du critère ES pour les 6 algorithmes testés, pour un bruit additif blanc, en fonction du type de signal et de la puissance du bruit (Faible : $\sigma = 0.5$, Moyen : $\sigma = 1$, Fort : $\sigma = 2$).

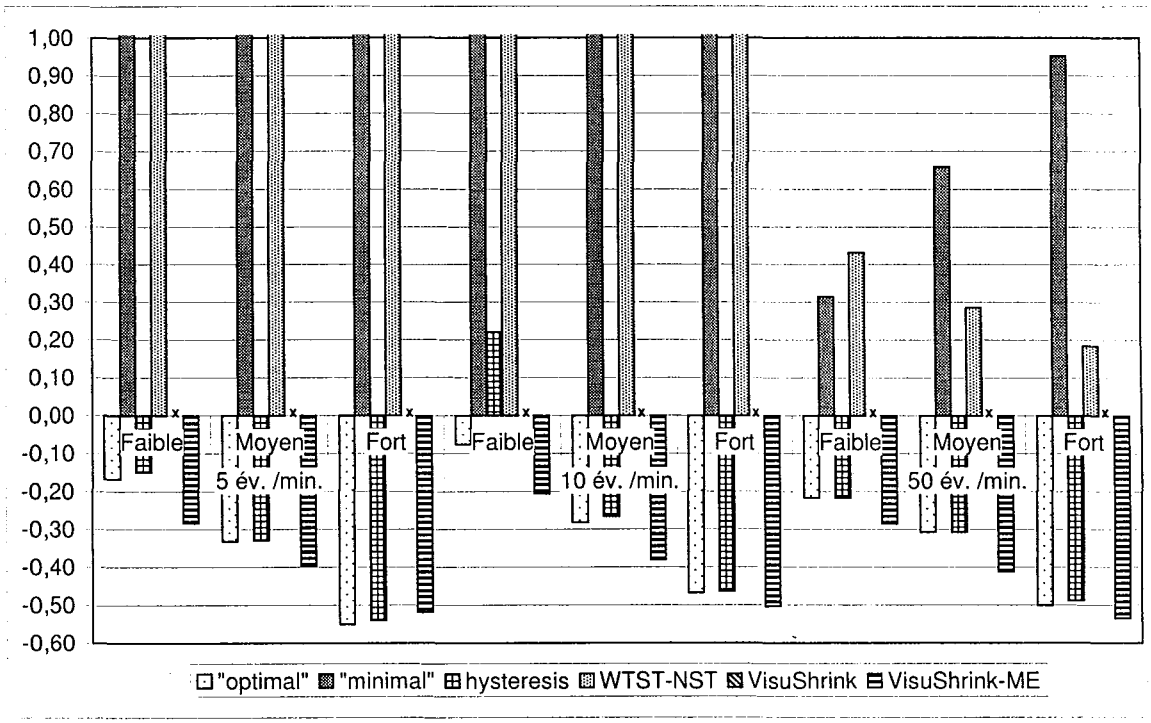


FIG. B.8 – Valeurs détaillées du critère ES pour les 6 algorithmes testés, pour un bruit additif coloré, en fonction du type de signal et de la puissance du bruit (Faible : $\sigma = 0.5$, Moyen : $\sigma = 1$, Fort : $\sigma = 2$).

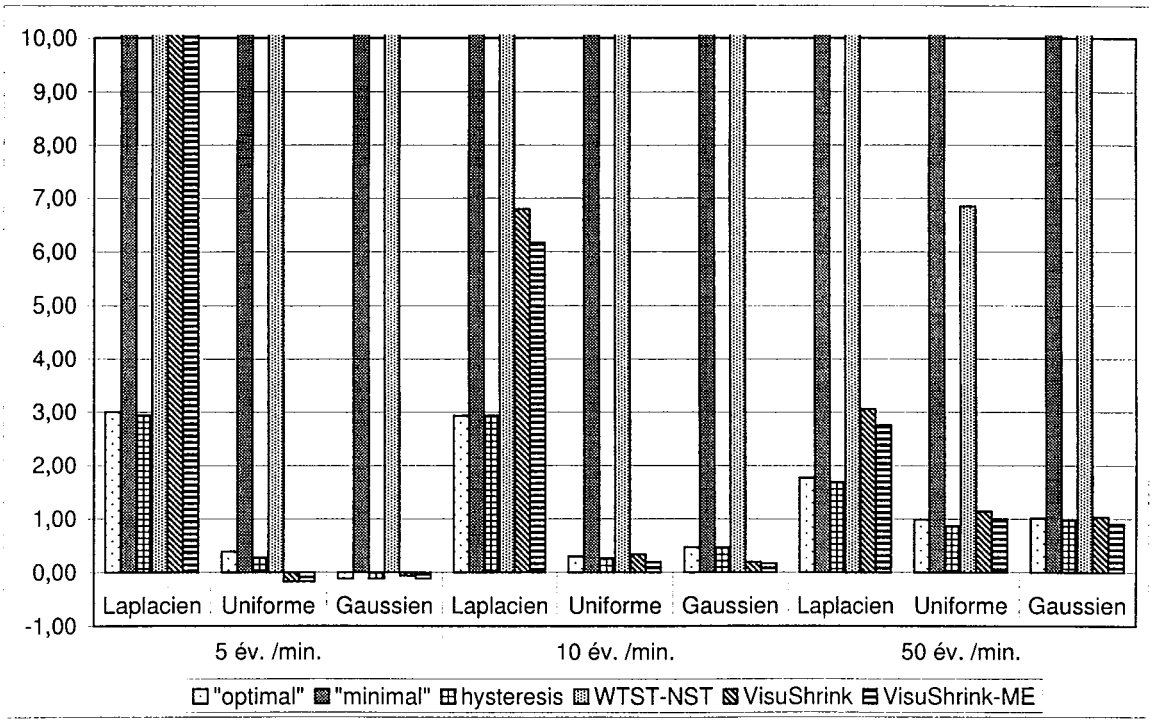


FIG. B.9 – Valeurs détaillées du critère *ED* pour les 6 algorithmes testés, pour un bruit additif blanc, en fonction du type de signal et de la distribution du bruit.

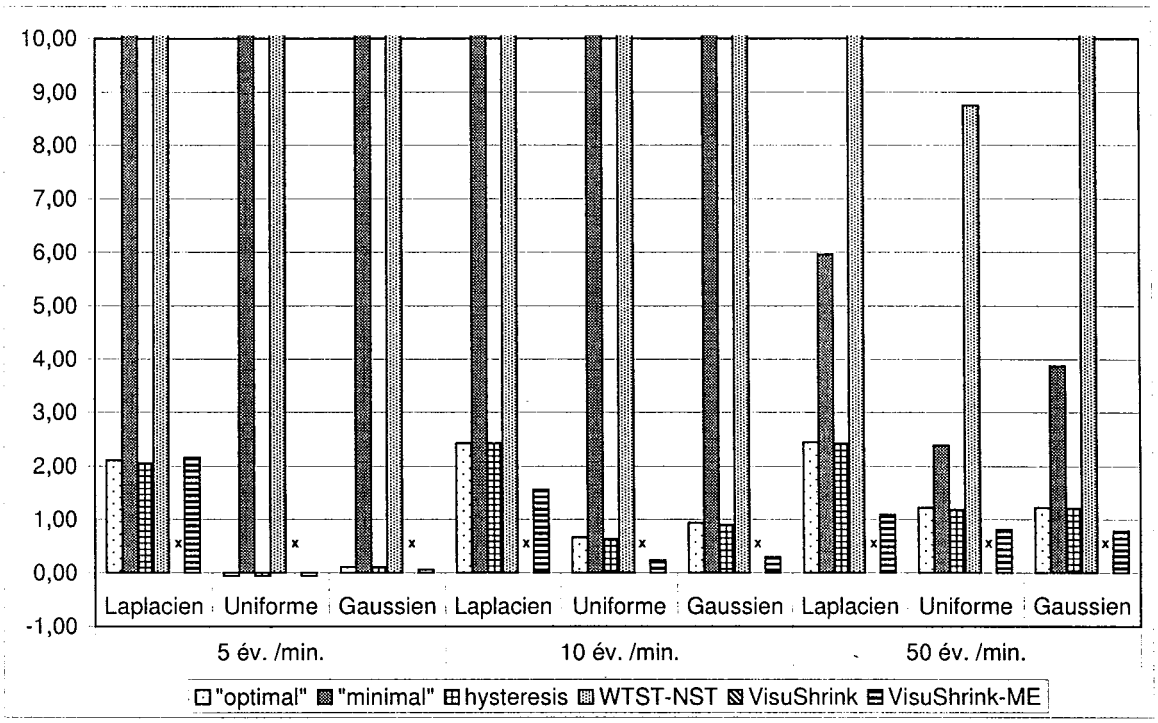


FIG. B.10 – Valeurs détaillées du critère *ED* pour les 6 algorithmes testés, pour un bruit additif coloré, en fonction du type de signal et de la distribution du bruit. L'algorithme *VisuShrink* n'est évidemment pas adapté au bruit coloré et il ne segmente pas le signal.

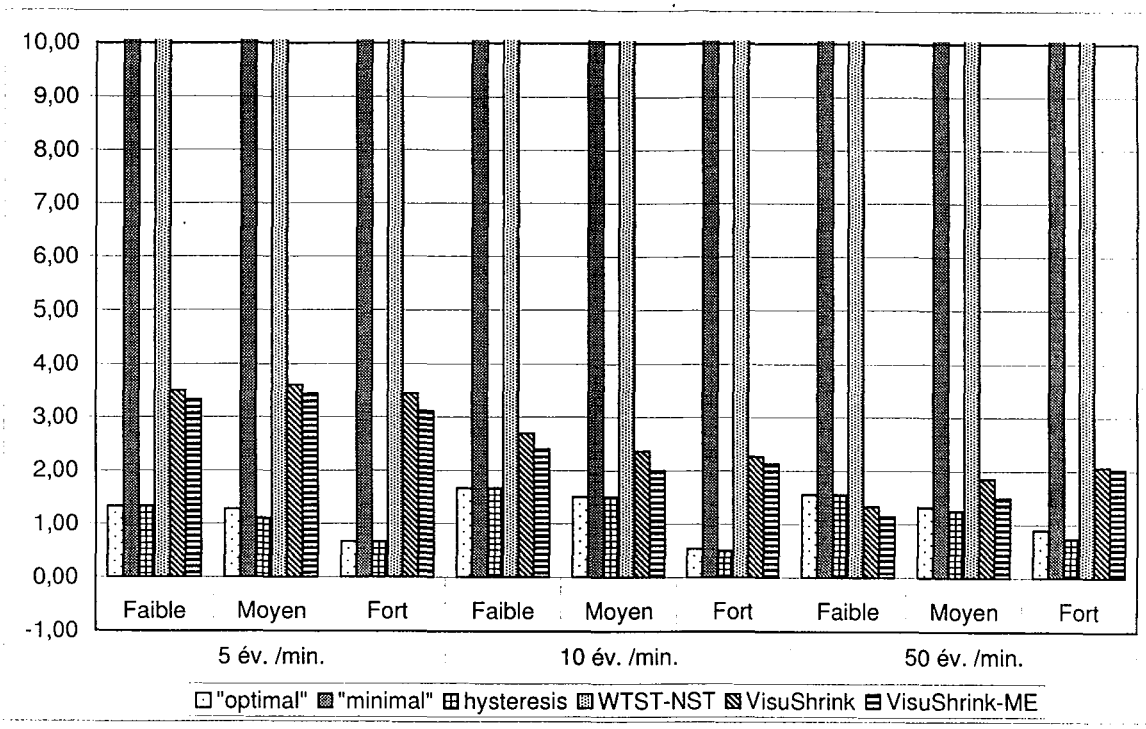


FIG. B.11 – Valeurs détaillées du critère *ED* pour les 6 algorithmes testés, pour un bruit additif blanc, en fonction du type de signal et de la puissance du bruit (Faible : $\sigma = 0.5$, Moyen : $\sigma = 1$, Fort : $\sigma = 2$).

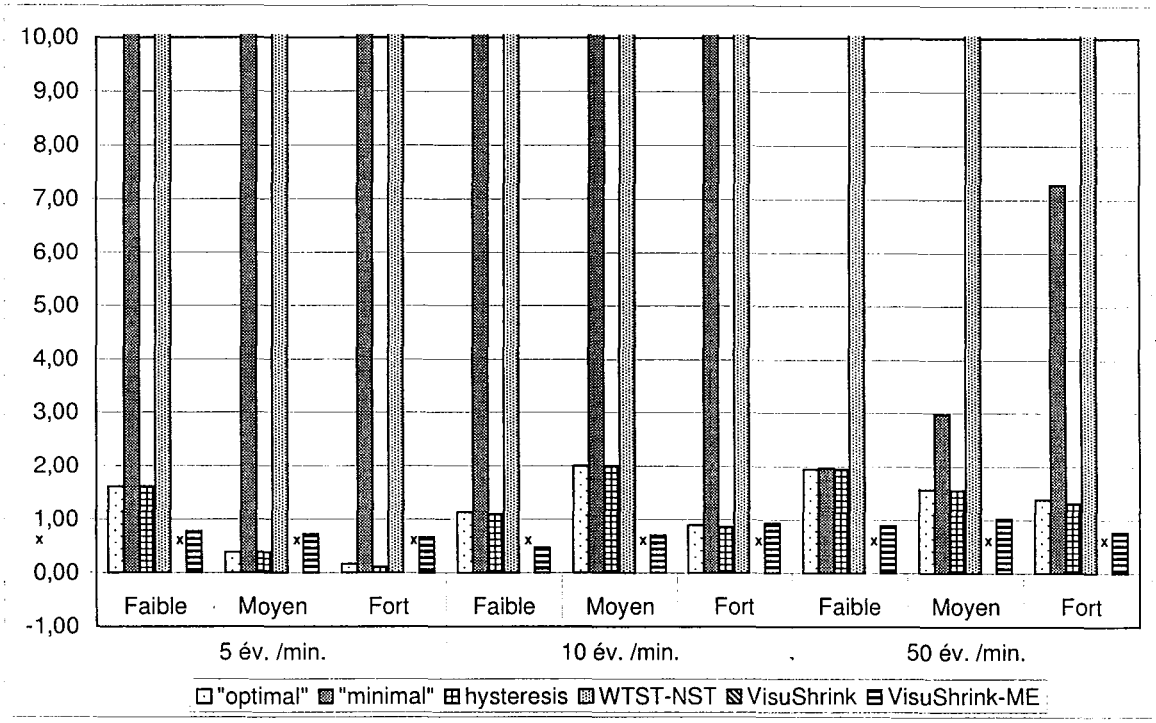


FIG. B.12 – Valeurs détaillées du critère *ED* pour les 6 algorithmes testés, pour un bruit additif coloré, en fonction du type de signal et de la puissance du bruit (Faible : $\sigma = 0.5$, Moyen : $\sigma = 1$, Fort : $\sigma = 2$). L'algorithme *VisuShrink* n'est évidemment pas adapté au bruit coloré et il ne segmente pas le signal.

Bibliographie

[1] * * *. *Statistics toolbox for with Matlab. Users guide, version 2.* The Mathworks, Inc., 1999.

[2] * * *. Interpreting abnormal abdominal sounds. *Nursing*, 30(6) :28, 2000.

[3] * * *. *Pathologie Fonctionnelle Digestive.* Laboratoire de Physiologie Digestive, Hôpital Broussais, Paris, 2002. `http://pfd.aphp.org/index.html`.

[4] A. AKIN et H. SUN. Time-frequency methods for detecting spike activity of stomach. *Med. Biol. Eng. Comput.*, 37(3) :381–390, 1999.

[5] T.J ANGTUACO, J. EVANS, L. HARRELL, H.A. MANSY, et R.H. SANDLER. The effect of abdominal palpation on gastrointestinal (GIS) in healthy subjects : A quantitative investigation using computerized analysis. *Am. J. Gastroenterology*, 95(9) :2491, 2000.

[6] A. ANTONIADIS, J. BIGOT, et T. SAPATINAS. Wavelet estimators in nonparametric regression : a comparative simulation study. *Journal of Statistical Software*, 6(6) :1–83, 2001. `http://www-lmc.imag.fr/lmc-sms/Anestis.Antoniadis/HTTP/publis-anto.html`.

[7] E. ARNBJÖRNSSON. Normal and pathological bowel sound patterns. *Ann. Chir. Gynaecol.*, 75(6) :314–318, 1986.

[8] F. AUGER, P. FLANDRIN, P. GONÇALVÈS, et O. LEMOINE. *Time-frequency toolbox for use with Matlab. Tutorial.* CNRS (France), Rice University (USA), 1995–1996.

[9] L. AUSTGEN, R.A. BOWEN, et M. ROUGE. *Pathophysiology of the digestive system.* Colorado State University, USA, 2002. `http://arbl.cvmbs.colostate.edu/hbooks/pathphys/digestion/index.html`.

[10] P. BERKHIN. Survey of clustering data mining techniques. Rapport technique, Accrue Software, San Jose, CA, 2002. `http://citeseer.nj.nec.com/berkhin02survey.html`.

[11] S.M. BERMAN. *Sojourns and Extremes of Stochastic Processes.* Wadsworth, Reading, MA, 1989.

[12] A. BIJAOUI. Wavelets, Gaussian mixture and Wiener filtering. *Signal Processing*, 82(4) :709–712, 1996.

[13] A. BLUM et P. LANGLEY. Selection of relevant features and examples in machine learning. *Artificial Intelligence*, 97(1–2) :245–271, 1997.

[14] S. BOUILLOT. Classification des bruits parasites dans l'habitacle d'un véhicule. Rapport technique, 01dB Acoustics & Vibration – Metravib Technologies, Limonest, 2003. Rapport de stage confidentiel.

[15] D. BRAY, R. REILLY, L. HASKIN, et B. MCCORMACK. Assessing motility through abdominal sound monitoring. Dans *Proceedings of IEEE-EMBS Conference*, pages 2398–2400. Chicago, USA, octobre 1997.

[16] D. BRIE. *Méthodes statistiques de détection de rupture de modèle. Application au traitement du signal issu d'un capteur à courants de Foucault.* Doctorat de l'université, spécialité automatique, Université de Nancy I, France, 1992.

[17] F. CAMPBELL, B. STOREY, P. CULLEN, et A. CUSCHIERI. Surface vibration analysis (SVA) : a new non-invasive monitor of gastrointestinal motility. *Gut*, 30(1) :39–45, 1989.

[18] W. CANNON. Auscultation of the rythmic sounds produced by the stomach and intestines. *Am. J. Physiology*, 14 :339–353, 1905.

[19] D.B CARR, C.J. YOUNG, R.C. ASTER, et X. ZHANG. Cluster analysis for CTBT seismic event monitoring. Dans *21st Seismic Research Symposium : Technologies for Monitoring the Comprehensive Nuclear-Test-Ban Treaty*, septembre 1999. `http://www.ctbt.rnd.doe.gov/Symposium1999/papers7/carr.pdf`.

[20] G. CELEUX. *Reconnaissance de mélanges de densités de probabilité et applications à la validation des résultats en classification.* Thèse de Doctorat d'État es Sciences Mathématiques, L'Université Paris IX Dauphine, 1987.

[21] G. CELEUX. Analyse discriminante. Dans G. GOVAERT, éditeur, *L'analyse des données*, pages 201–234. Hermès, 2003.

[22] J. CHEN, Z. LIN, Q. WU, et R. MCCALLUM. Non-invasive identification of gastric contractions from surface electrogastrogram using back-propagation neural networks. *Med. Eng. Phys.*, 17(3) :219–225, 1995.

[23] J. CHEN, R. MCCALLUM, et R. RICHARDS. Frequency components of the electrogastrogram and their correlations with gastrointestinal contractions in humans. *Med. Biol. Eng. Comput.*, 31(1) :60–67, 1993.

[24] P-A. CHÉVALIER. Une introduction didactique à la géometrie du GPS. Ecole d'ingénieurs de Bienne, Haute Ecole Spécialisée Bernoise, 2002. `http://www.hta-bi.bfh.ch/~cip/gps/ressources/partie2.pdf`.

[25] R. COIFMAN et D. DONOHO. Translation-invariant de-noising. Dans A. ANTONIADIS et G. OPPENHEIM, éditeurs, *Wavelets and Statistics*, pages 125–150. Springer-Verlag, 1995. `http://citeseer.nj.nec.com/coifman95translationinvariant.html`.

[26] R. COIFMAN et Y. MEYER. Remarques sur l'analyse de Fourier à fenêtre. *C. R. Académie Sciences Paris*, 312 :259–261, 1991.

[27] R. COIFMAN et M. WICKERHAUSER. Entropy-based algorithms for best basis selection. *IEEE Trans. Information Theory*, 38 :712–718, 1992.

[28] R. COIFMAN et M. WICKERHAUSER. Adapted waveform de-noising for medical signals and images. *IEEE Engineering in Medicine and Biology Magazine*, 14(5) :578–586, 1995.

[29] R. COIFMAN et M. WICKERHAUSER. Experiments with adapted wavelet de-noising for medical signals and images. Dans M. AKAY, éditeur, *Time-Frequency and Wavelets in Biomedical Engineering*, pages 323–346. IEEE Press, 1998.

[30] B. CRAINE, M. SILPA, et C. O'TOOLE. Computerized auscultation applied to irritable bowel syndrome. *Digestive Diseases and Sciences*, 44(9) :1887–1892, 1999.

[31] B. CRAINE, M. SILPA, et C. O'TOOLE. Enterotachogram analysis to distinguish irritable bowel syndrome from Crohn's disease. *Digestive Diseases and Sciences*, 46(9) :1974–1979, 2001.

[32] B. CRAINE, M. SILPA, et C. O'TOOLE. Two-dimensional positional mapping of gastrointestinal sounds in control and functional bowel syndrome patients. *Digestive Diseases and Sciences*, 47(6) :1290–1296, 2002.

[33] P. CULLEN, B. STOREY, A. CUSCHIERI, et F. CAMPBELL. Detection of clustered gastrointestinal contractions in partial intestinal obstruction by surface vibration analysis. *Annals of Surgery*, 210(2) :234–238, 1989.

[34] D. DALLE, G. DEVROEDE, R. THIBAULT, et J. PERRAULT. Computer analysis of bowel sounds. *Comput. Biol. Med.*, 4 :247–256, 1975.

[35] M. DASH et H. LIU. Feature extraction for classification. *Intelligent Data Analysis*, 1(1–4) :131–156, 1997.

[36] M. DASH, H. LIU, et J. YAO. Dimensionality reduction of unsupervised data. Dans *Tools with Artificial Intelligence, 1997. Proceedings., Ninth IEEE International Conference on*, numéro 3-8, pages 532–539, 1997. `http://citeseer.nj.nec.com/liu97dimensionality.html`.

[37] D. DONOHO. De-noising via soft thresholding. Rapport technique, Department of Statistics, Stanford University, 1992. `http://www.stat.stanford.edu/~donoho/Reports/index.html`.

[38] D. DONOHO et I. JOHNSTONE. Ideal spatial adaptation via wavelet shrinkage. *Biometrika*, 81 :425–455, 1994.

[39] D.J. DU PLESSIS. Clinical observation on intestinal motility. *South African Medical J.*, 28 :27–33, 1954.

[40] R.O. DUDA, P.E. HART, et D.G. STORK. *Pattern classification.* John Wiley & sons, inc., deuxième édition, 2001.

[41] J. FARRAR et F. INGELFINGER. Gastrointestinal motility as revealed by study of abdominal sounds (with discussion). *Gastroenterology*, 29(5) :789–802, 1955.

[42] D. FASULO. An analysis of recent work on clustering algorithms, 1999. `http://citeseer.nj.nec.com/fasulo99analysi.html`.

[43] P. FLANDRIN. *Temps-fréquence.* Hermes, deuxième édition, 1998.

[44] W. FOULK, C. CODE, C. MORLOCK, et J. BARGEN. A study of the motility patterns and the basic rhythm in the duodenum and upper part of the jejunum of human beings. *Gastroenterology*, 26(4) :601–611, 1954.

[45] C. FRALEY et A.E. RAFTERY. How many clusters ? Which clustering method ? Answers via model-based cluster analysis. *The Computer Journal*, 41(8) :578–588, 1998. `http://citeseer.nj.nec.com/fraley98how.html`.

[46] R.A. FREITAS JR.. *Nanomedecine.* 2003. http://www.nanomedicine.com/.

[47] J. GADE, P. KRUSE, O. TRIER ANDERSEN, S. BOEL PEDERSEN, et S. BOESBY. Physicians abdominal auscultation – a multi-rater agreement study. *Scandinavian Journal of Gastroenterology*, 33(7) :773–777, 1998.

[48] C. GARNER et H. EHRENREICH. Non invasive topographic analysis of intestinal activity in man on the basis of acoustic phenomena. *Res. Exp. Med. (Berl.)*, 189(2) :129–140, 1989.

[49] M. GAUSSEN. L'auscultation de l'abdomen. *Journal de Médecine de Bordeaux*, 1 :25–31, 1952.

[50] D. GERHARD. Audio signal classification : an overview. Canadian Artificial Intelligence, 2000. http://www2.cs.uregina.ca/~gerhard/publications.html.

[51] P. GIMONDO et P. MIRK. A new method for evaluating small intestinal motility using duplex Doppler sonography. *Am. J. Roentgenology*, 168(1) :187–192, 1997.

[52] D. GOUNOT. Ultrasons et échographie i : physique des ultrasons (cours). Institut de Physique Biologique, Strasbourg, 2001. http://www-ipb.u-strasbg.fr/ipb/edu/pcem2/echo1/index.htm.

[53] N. GRIP. *Hilbert space frames and bases, a comparison of Gabor and wavelet frames and application to multicarrier digital communications.* Licentiate thesis, Department of Mathematics, Lulea University of Technology, Suède, 2000.

[54] L. HADJILEONTIADIS, L. LIATSOS, C. MAVROGIANNIS, T. ROKKAS, et S. PANAS. Enhancement of bowel sounds by wavelet-based filtering. *IEEE Trans. Biomedical Engineering*, 47(7) :876–886, 2000.

[55] L. HADJILEONTIADIS et S. PANAS. Separation of discontinuous adventitious sounds from vesicular sounds using a wavelet-based filter. *IEEE Trans. Biomedical Engineering*, 44(12) :1269–1281, 1997.

[56] P. HELLSTROM. Motility of small intestine : a case for pattern recognition. *Journal of Internal Medicine*, 237(4) :391–394, 1995.

[57] N. HESS-NIELSEN et M. WICKERHAUSER. Wavelets and time-frequency analysis. *Proceedings of the IEEE. Special issue on wavelet applications*, 84(4) :523–540, 1996. http://www.math.wustl.edu/~victor/papers/wtfa.pdf.

[58] S. HOLT, W. MCDICKEN, T. ANDERSON, I. STEWART, et R. HEADING. Dynamic imaging of the stomach by real-time ultrasound – a method for the study of gastric motility. *Gut*, 21(7) :597–601, 1980.

[59] A.K. JAIN, M.N. MURTY, et P.J. FLYNN. Data clustering : a review. *ACM Computing Surveys*, 31(3) :264–323, 1999. http://citeseer.nj.nec.com/jain99data.html.

[60] C. JUTTEN *et al.* . ELENA – Enhanced learning for evolutive neural architecture. Rapport technique, Esprit III Project, Institut National Polytechnique de Grenoble, 1995. http://www.ph.tn.tudelft.nl/PRInfo/data/msg00005.html.

[61] H.M. KALAYEH et D.A. LANDGREBE. Predicting the required number of training samples. *IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI*, 5(6) :664–667, 1983.

[62] J. KELLOW, T. BORODY, S. PHILIPPS, R. TUCKER, et A. HADDAD. Human interdigestive motility : variations in patterns from oesophagus to colon. *Gastroenterology*, 91(2) :386–395, 1986.

[63] M. KHALIL et J. DUCHÊNE. Detection and classification of multiple events in piecewise stationary signals : Comparison between autoregressive and multiscale approaches. *Signal Processing*, 75(3) :239–251, 1999.

[64] E. KOLATCH. Clustering algorithms for spatial databases : A survey. `http://citeseer.nj.nec.com/436843.html`.

[65] D. LANDGREBE. Information extraction principles and methods for multispectral and hyperspectral image data. Dans C.H. CHEN, éditeur, *Information Processing for Remote Sensing*. World Scientific Publishing Co., Inc., 2000. `http://dynamo.ecn.purdue.edu/~landgreb/Principles.pdf`.

[66] D. LI, K.D WONG, Y.H. HU, et A.M. SAYEED. Detection, classification and tracking of targets. *IEEE Signal Processing Magazine*, 19(2) :17–29, 2002.

[67] J. LIANG et J. CHEN. What can be measured from surface electrogastrography – computer simulations. *Digestive Diseases and Sciences*, 42(7) :1331–1343, 1997.

[68] G. LINDBERG, M. IWARZON, P. STÅL, et R. SEENSALU. Digital ambulatory monitoring of small-bowel motility. *Scandinavian Journal of Gastroenterology*, 25 :216–224, 1990.

[69] W.Y. LIU, I.E. MAGNIN, et G. GIMENEZ. Un nouvel opérateur pour la détection de ruptures dans des signaux bruités. *Traitement du Signal*, 12(3) :225–235, 1995.

[70] L. LJUNG et T. GLAD. *Modeling of dynamic systems*. Prentice Hall, 1994.

[71] S. MALLAT. *A wavelet tour of signal processing*. Academic Press, 1999.

[72] H. MANSY et R. SANDLER. Bowel sound signal enhancement using adaptive filtering. *IEEE Engineering in Medicine and Biology Magazine*, 16(6) :105–117, 1997.

[73] H. MANSY et R. SANDLER. Choice of operating parameters in heart sound removal from bowel sounds using adaptive filtering. Dans *Proc. IEEE EMBS*, pages 1398–1401, 1997.

[74] H. MANSY et R. SANDLER. Detection and analysis of gastrointestinal sounds in normal and small bowel obstructed rats. *Med. Biol. Eng. Comput.*, 38(1) :42–48, 2000.

[75] H.A. MANSY, T.J. ROYSTON, et R.H. SANDLER. Acoustic characteristics of air cavities at low audible frequencies with application to pneumoperitoneum detection. *Med. Biol. Eng. Comput.*, 39 :159–167, 2001.

[76] G.W. MILTON. Normal bowel sounds. *Med. J. Aust.*, 2 :490–493, 1958.

[77] M. MISITI, Y. MISITI, G. OPPENHEIM, et J-M. POGGI. *Lès ondelettes et leurs applications*. Hermes, 2003.

[107] J.-L. STARCK et A. BIJAOUI. Filtering and deconvolution by the wavelet transform. *Signal Processing*, 35 :195–211, 1994.

[108] J.-L. STARCK et F. MURTAGH. Image restoration using wavelet transform. *Astronomy and Astrophysics*, 288 :342–348, 1994.

[109] M-C. SU. A new index of cluster validity. `http://selab.csie.ncu.edu.tw/~muchun/course/cluster/A%20New%20Index%20of%20Cluster%20Validity.pdf`.

[110] M. SUGRUE et M. REDFERN. Computerized phonoenterography : the clinical evaluation of a new system. *J. Clin. Gastroenterol.*, 18(2) :139–144, 1994.

[111] A. SUZUKI, C. SUMI, K. NAKAYAMA, et M. MORI. Real-time adaptive cancelling of ambient noise in lung sound measurement. *Med. Biol. Eng. Comput.*, 33 :704–708, 1995.

[112] A. TESEI et C. REGAZZONI. HOS-based generalized noise PDF models for signal detection optimization. *Signal Processing*, 65(2) :267–281, 1998.

[113] M. THIRIET. *Appareil digestif.* INRIA, Projet M3N, 2000. `http://www-rocq1.inria.fr/Marc.Thiriet/Glosr/Bio/TubDigest/index.html`.

[114] A.B.R. THOMSON et E.A. SHAFFER. *Principes fondamentaux de gastro-entérologie.* Gastroenterology resource center, troisième édition, 2001. `http://gastroresource.com`.

[115] H. THORNBURG. Detection and modeling of transients. Center for Computer Research in Music and Acoustics (CCRMA), Department of Music, Stanford University, California, 2002. `http://www-ccrma.stanford.edu/~jos/tvar/tvar.html`.

[116] T. TOMOMASA, A. MORIKAWA, R. SANDLER, H. MANSY, H. KONEKO, T. MASAHIKO, P. HYMAN, et Z. ITOH. Gastrointestinal sounds and migrating motor complex in fasted humans. *Am. J. Roentgenology*, 94(2) :374–381, 1999.

[117] T. TOMOMASA, A. MORIKAWA, M. TABATA, P. HYMAN, et Z. ITOH. Bowel sounds reflect gastrointestinal motility and small intestinal transit time in fasted humans. *Gastroenterology*, 112(4) :A839, 1997.

[118] M. UNSER et A. ALDROUBI. A review of wavelets in biomedical applications. *Proc. IEEE*, 84(4) :626–638, 1996.

[119] C. VASSEUR, G. DEVROEDE, D. DALLE, N. VAN HOUTTE, E. BASTIN, et R. THIBAULT. Postprandial bowel sounds. *IEEE Trans. Biomedical Engineering*, BME–22(5) :443–448, 1975.

[120] B. WALDRON, B. STOREY, D. SMITH, P. CULLEN, et F. CAMPBELL. Computerised method for acquisition and display of gastrointestinal motility data. *Med. Biol. Eng. Comput.*, 29(3) :304–308, 1991.

[121] W. WATSON et E. KNOX. Phonoenterography : the recording and analysis of bowel sounds. *Gut*, 8(1) :88–94, 1967.

[122] C. WELLS, L. TINCKLER, K. RAWLINSON, et H. JONES. Postoperative gastrointestinal motility. *The Lancet*, 1 :4–10, 1964.

[123] F. WENDLING. *Mise en correspondance d'observations EEG de profondeur pour la reconnaissance de signatures spatio-temporelles dans les crises d'épilepsie.* Doctorat de l'université, mention traitement du signal et télécommunications, Université de Rennes I, France, 1996.

[124] M. WICKERHAUSER. INRIA lectures on wavelet packet algorithms. Dans Pierre-Louis LIONS, éditeur, *Problèmes Non-Linéaires Appliqués, Ondelettes et Paquets D'Ondes*, pages 31–99. INRIA, Rocquencourt, France, 17–21 juin 1991. Minicourse lecture notes, `http://www.math.wustl.edu/~victor/papers/lwpa.pdf`.

[125] B. WIDROW, J. GLOVER, J. MCCOOL, J. KAUNITZ, C. WILLIAMS, R. HEARN, J. ZEIDLER. E. DONG, et R. GOODLIN. Adaptive noise cancelling : principles and applications. *Proc. IEEE*, 63(12) :1692–1716, 1975.

[126] H. YOSHINO, Y. ABE, T. YOSHINO, et K. OHSATO. Clinical application of spectral analysis of bowel sounds in intestinal obstruction. *Dis. Col. Rectum*, 33(9) :753–757, 1990.

[127] M. YUKI, K. ADACHI, H. FUJISHIRO, Y. UCHIDA, Y. MIYAOKA, N. YOSHINO, T. YUKI, M. ONO, et Y. KINOSHITA. Is a computerized bowel sound auscultation useful for the detection of increased bowel motility ? *Am. J. Gastroenterology*, 97(7) :1846–1848, 2002.

[128] Y. ZHENG, D.B.H. TAY, et L. LI. Signal extraction and power spectrum estimation using wavelet transform scale space filtering and Bayes shrinkage. *Signal Processing*, 80(8) :1535–1549, 2000.

Résumé

L'objectif de ce travail de recherche est le développement d'un système d'étude de sons, plus particulièrement dédié à la phonoentérographie, qui devrait aboutir à plus long terme à un outil d'aide au diagnostic. La première étape présente une chaîne d'instrumentation multi-voies spécifique. Elle est suivie par le pré-traitement : la détection, la segmentation et le débruitage par ondelettes sont réalisés avec un algorithme original optimisé par une méthode de point-fixe. Une deuxième phase introduit des connaissances a priori sur les sons abdominaux et étudie leur localisation spatiale. Les caractéristiques physiques (fréquence, intensité, durée) décrivent les sons individuellement. L'étude globale des phonoentérogrammes est réalisée à partir d'indices d'activité (nombre d'événements, énergie moyenne, etc.). Les caractéristiques physiques et les indices sont utilisés dans l'analyse statistique des signaux, par analyse en composantes principales et classification non supervisée.

Mots-clés: Phonoentérographie, aide au diagnostic, signaux non-stationnaires, détection par ondelettes, débruitage par ondelettes, localisation, extraction de caractéristiques, analyse en composantes principales, classification.

Abstract

The goal of this research is the development of an abdominal sound analysis system. The long-term objective is a diagnostic aid system based upon phonoenterogram analysis. The first step is the study of a specific multi-channel acquisition system. The following stage is the preprocessing : we implement an original wavelet fixed-point algorithm for the abdominal sounds detection, segmentation and denoising. A second preprocessing step introduces available a priori knowledge on physical characteristics of isolated abdominal sounds (main frequency, duration, acoustic intensity) and studies the problem of sound source localization. The extracted physical features, as well as a set of activity indexes appropriated for the global description of the phonoenterograms (number of sounds, mean energy, etc.), are used in the last processing stage, the signal analysis. The employed techniques are principal component analysis and unsupervised clustering.

Keywords: Phonoenterography, diagnosis aid, non-stationary signals, wavelet detection and denoisng, localization, feature extraction, principal component analysis, clustering.

INSTITUT NATIONAL
POLYTECHNIQUE
DE LORRAINE

AUTORISATION DE SOUTENANCE DE THESE DU DOCTORAT DE L'INSTITUT NATIONAL POLYTECHNIQUE DE LORRAINE

o0o

VU LES RAPPORTS ETABLIS PAR :
Madame Régine LE BOUQUIN JEANNES, Professeur, LTSI, Université de Rennes 1
Monsieur Gérard GIMENEZ, Professeur, Secteur Administration, Directeur CERMEP, Lyon

Le Président de l'Institut National Polytechnique de Lorraine, autorise :

Monsieur RANTA Serban Radu

à soutenir devant un jury de l'INSTITUT NATIONAL POLYTECHNIQUE DE LORRAINE, une thèse intitulée :

" Traitement et analyse de signaux sonores physiologiques : application à la phonoentérographie "

NANCY BRABOIS
2, AVENUE DE LA
FORET-DE-HAYE
BOITE POSTALE 3
F - 54501
VANŒUVRE CEDEX

en vue de l'obtention du titre de :

DOCTEUR DE L'INSTITUT NATIONAL POLYTECHNIQUE DE LORRAINE

Spécialité : « **Automatique et Traitement du Signal** »

Service Commun de la Documentation
INPL
Nancy-Brabois

Fait à Vandoeuvre, le 26 novembre 2003
Le Président de l'I.N.P.L.,
L. SCHUFFENECKER

INSTITUT NATIONAL POLYTECHNIQUE DE LORRAINE - NANCY -

Pour le Président
Le Vice-Président,

J.-L. COUJARD

TEL. 33/03.83.59.59.59
FAX. 33/03.83.59.59.55

Résumé

L'objectif de ce travail de recherche est le développement d'un système d'étude de sons, plus particulièrement dédié à la phonoentérographie, qui devrait aboutir à plus long terme à un outil d'aide au diagnostic. La première étape présente une chaîne d'instrumentation multi-voies spécifique. Elle est suivie par le pré-traitement : la détection, la segmentation et le débruitage par ondelettes sont réalisés avec un algorithme original optimisé par une méthode de point-fixe. Une deuxième phase introduit des connaissances a priori sur les sons abdominaux et étudie leur localisation spatiale. Les caractéristiques physiques (fréquence, intensité, durée) décrivent les sons individuellement. L'étude globale des phonoentérogrammes est réalisée à partir d'indices d'activité (nombre d'événements, énergie moyenne, etc.). Les caractéristiques physiques et les indices sont utilisés dans l'analyse statistique des signaux, par analyse en composantes principales et classification non supervisée.

Mots-clés: Phonoentérographie, aide au diagnostic, signaux non-stationnaires, détection par ondelettes, débruitage par ondelettes, localisation, extraction de caractéristiques, analyse en composantes principales, classification.

Abstract

The goal of this research is the development of an abdominal sound analysis system. The long-term objective is a diagnostic aid system based upon phonoenterogram analysis. The first step is the study of a specific multi-channel acquisition system. The following stage is the preprocessing : we implement an original wavelet fixed-point algorithm for the abdominal sounds detection, segmentation and denoising. A second preprocessing step introduces available a priori knowledge on physical characteristics of isolated abdominal sounds (main frequency, duration, acoustic intensity) and studies the problem of sound source localization. The extracted physical features, as well as a set of activity indexes appropriated for the global description of the phonoenterograms (number of sounds, mean energy, etc.), are used in the last processing stage, the signal analysis. The employed techniques are principal component analysis and unsupervised clustering.

Keywords: Phonoenterography, diagnosis aid, non-stationary signals, wavelet detection and denoisng, localization, feature extraction, principal component analysis, clustering.