

TABLE DES MATIÈRES

RÉSUMÉ	iii
LISTE DES FIGURES	vii
LISTE DES TABLEAUX	ix
LISTE DES ABRÉVIATIONS	x
CHAPITRE I	
DÉFINITIONS ET CLASSIFICATION DES SONS RESPIRATOIRES	3
1.1 Nomenclature et classification des sons respiratoires	3
1.1.1 Sons normaux	4
1.1.2 Sons adventices continus	7
1.1.3 Sons adventices discontinus	8
1.2 Base de données utilisée	10
1.3 Problématique de recherche	11
1.4 Objectifs	12
1.4.1 Analyse des crépitants	12
1.4.2 Analyse des sibilants	14
CHAPITRE II	
EXTRACTION DES CRÉPITANTS DES SONS VÉSICULAIRES	15
2.1 Principe de séparation stationnaire et non-stationnaire des sons respiratoires	15
2.2 Extraction des crépitants par filtre non-linéaire	16
2.2.1 Filtre prédictif	16
2.2.2 Filtre de séparation non-linéaire ST-NST	17
2.3 Extraction des crépitants par logique floue	20
2.3.1 Concept de base de la logique floue	20
2.3.2 Système d'inférence floue (FIS)	25

2.3.3	Système d'inférence floue à base d'un réseau adaptatif (ANFIS) . . .	27
2.3.4	Filtre de séparation basé sur la logique floue GFST-NST	29
2.4	Extraction des crépitants par la transformée en ondelettes	30
2.4.1	Transformée en ondelettes	30
2.4.2	Analyse multirésolution	31
2.4.3	Filtre de séparation basé sur la transformée en ondelettes WTST-NST	34
2.5	Extraction des crépitants par les paquets d'ondelettes	35
2.5.1	Transformée par paquets d'ondelettes	37
2.5.2	Recherche de la meilleure base	38
2.5.3	Filtre de séparation basé sur les paquets en ondelettes WPST-NST	39
CHAPITRE III		
IDENTIFICATION DES CRÉPITANTS 43		
3.1	Principe	43
3.2	Filtre de débruitage	44
3.3	Détection des crépitants	45
3.3.1	Choix du paramètre de la détection	46
3.3.2	Notion de la dimension fractale	47
3.3.3	Détection par le paramètre "Dimension Fractale"	50
3.4	Classification des crépitants	52
3.4.1	Caractérisation temps-fréquence	54
3.4.2	Modèle de Prony	55
3.4.3	Transformée en ondelettes continue	58
CHAPITRE IV		
DÉTECTION DES SIBILANTS 62		
4.1	Principe	62
4.2	Calcul du spectrogramme	63
4.3	Détection des sibilants par le Laplacien d'une gaussienne	64
4.3.1	Laplacien d'une gaussienne (LoG)	65

4.3.2	Détection des sibilants par filtre LoG	66
4.4	Détection des sibilants par modélisation auditive	68
4.4.1	Système auditif	68
4.4.2	Description de l'algorithme (FDDT)	69
4.5	Détecteur d'épisodes des sibilants (WED)	73
4.6	Représentation des sibilants par "Chirplets"	76
4.6.1	Transformée par "Chirplets"	76
4.6.2	Représentation des sibilants par "Chirplets"	78
CHAPITRE V		
RÉSULTATS EXPÉRIMENTAUX 80		
5.1	Préparation de la base de données	80
5.2	Critères d'évaluation	82
5.2.1	Taux de séparation	82
5.2.2	Rapport signal-sur-bruit (SNR)	82
5.2.3	Sensitivité et valeur prédictive positive	83
5.2.4	Performance de la classification	83
5.3	Résultats de la séparation des crépitants	83
5.4	Résultats de la détection des crépitants	90
5.5	Résultats de la classification des crépitants	91
5.6	Résultat de l'extraction des sibilants	93
CONCLUSION 97		
BIBLIOGRAPHIE 100		

LISTE DES FIGURES

1.1	Diagramme de classification des sons respiratoires.	5
1.2	Différents types de sons respiratoires normaux	5
1.3	Spectrogrammes de sons adventices continus	8
1.4	Spectrogrammes de sons adventices discontinus	9
1.5	Définitions de IDW, 2CD et PF	9
1.6	Schéma du système d'analyse des crépitants	13
2.1	Structure du filtre prédictif	17
2.2	Schéma du filtre ST-NST	18
2.3	Fonction non-linéaire utilisée par le filtre ST-NST	19
2.4	Diagramme d'un filtre numérique pour l'extraction des crépitants	20
2.5	Comparaison d'un ensemble classique et d'un ensemble flou	21
2.6	Définitions possibles de l'ensemble de "l'eau chaude".	23
2.7	Exemples de fonctions caractéristiques et de fonctions d'appartenance.	23
2.8	Trois exemples de fonctions d'appartenance.	23
2.9	Schéma du système d'inférence floue (FIS)	25
2.10	Types de règles floues SI-ALORS et les mécanismes de raisonnement flou	26
2.11	Système ANFIS utilisé dans le filtre GFST-NST	28
2.12	Schéma du filtre de séparation basé sur la logique floue GFST-NST	29
2.13	Schéma d'analyse MRD et de reconstitution MRR de l'algorithme de Mallat	32
2.14	Répartition fréquentielle par transformée en ondelettes.	33
2.15	Schéma du filtre de séparation basé sur les ondelettes WTST-NST	36
2.16	Répartition fréquentielle par paquets d'ondelettes.	37
2.17	Deux techniques de recherche de la meilleure base.	38
2.18	Exemple d'analyse de crépitants par paquets d'ondelettes.	39
2.19	Schéma du filtre de séparation basé sur les paquets en ondelettes WPST-NST	42
3.1	Comparaison de trois paramètres utilisés dans la détection des crépitants	47
3.2	Construction du triangle de <i>Sierpinski</i>	48

3.3	Dimension fractale calculée pour quatre objets	49
3.4	Algorithme automatique d'énumération des crépitants.	53
3.5	Exemple de détection des crépitants par dimension fractale	54
3.6	Caractérisation temps-fréquence d'un crépitant	56
3.7	Modélisation par modèle de Prony	57
3.8	La composante réelle de l'ondelette de Morlet	59
3.9	Crépitant simulé	60
3.10	Analyse par ondelettes d'un crépitant réel	61
4.1	Exemple de spectrogramme d'un son respiratoire avec sibilants	64
4.2	Filtrage d'images par le Laplacien d'une gaussienne (LoG)	66
4.3	Détection des sibilants en utilisant le Laplacien d'une gaussienne (LoG) . .	67
4.4	Schéma de la détection des sibilants par modélisation auditive (FDDT). . .	71
4.5	Définition des voisins pour l'algorithme CCL	72
4.6	Exemple de détection des sibilants par modélisation auditive	73
4.7	Schéma du détecteur d'épisodes de sibilants (WED).	74
4.8	Résultat de la détection par le détecteur d'épisodes du sibilant	76
4.9	Représentation du sibilant par chirplets.	78
5.1	Comparaison des performances des quatre filtres de séparation des crépitants	85
5.2	Performances des filtres WPST–NST et WTST–NST obtenus avec des crépitants fins et des gros crépitants	87
5.3	Performances des filtres WPST–NST et WTST–NST obtenus avec sons normaux et sibilants	89

LISTE DES TABLEAUX

1.1	Définitions et appellations des sons adventices	6
1.2	Caractéristiques des sons adventices et les pathologies associés	10
2.1	Degré d'appartenance à trois ensembles flous.	24
5.1	Base de données utilisée.	81
5.2	Paramètres du filtre WPST-NST et filtre WTST-NST.	86
5.3	Taux de séparation (TS) et temps de calcul (TR) des filtres de séparation WPST-NST et WTST-NST.	88
5.4	Niveau du bruit SNR introduit par les filtres de séparation WPST-NST et WTST-NST	90
5.5	Résultats de la détection par dimension fractale (algorithme de Sevcik) . .	92
5.6	Comparaison des performances des méthodes de classification des crépitants.	94
5.7	Performances des méthodes FDDT et WED d'extraction des sibilants . . .	96

LISTE DES ABRÉVIATIONS

<i>2CD</i>	: <i>Two Cycle Duration</i>	: Durée des deux premiers cycles
<i>ANFIS</i>	: <i>Adaptive-Network-based Fuzzy Inference System</i>	: Système d'inférence floue à base de réseau-adaptatif
<i>ATS</i>	: <i>American Thoracic Society</i>	: Association thoracique américaine
<i>EM</i>	: <i>Expectation-Maximisation</i>	: Estimation - maximisation
<i>ERB</i>	: <i>Equivalent Rectangular Bandwidth</i>	: Largeur de bande rectangulaire équivalente
<i>FD</i>	: <i>Fractal Dimension</i>	: Dimension fractale
<i>FIS</i>	: <i>Fuzzy Inference System</i>	: Système d'inférence floue
<i>GBW</i>	: <i>Gaussian Bandwidth</i>	: Largeur de la bande gaussienne
<i>GMM</i>	: <i>Gaussian Mixture Model</i>	: Modélisation par mélange de gaussiennes
<i>IDW</i>	: <i>Initial Deflexion Width</i>	: Largeur de la déflexion initiale
<i>LMS</i>	: <i>Least Mean Square</i>	: Minimum de la moyenne quadratique
<i>LoG</i>	: <i>Laplacian of Gaussian</i>	: Laplacien d'une gaussienne
<i>MDW</i>	: <i>Maximal Deflexion Width</i>	: Durée de la déflexion maximale
<i>MLE</i>	: <i>Maximum Likelihood Estimation</i>	: Estimation du maximum de vraisemblance
<i>MRD</i>	: <i>Multi-Resolution Decomposition</i>	: Décomposition par multirésolution
<i>MRR</i>	: <i>Multi-Resolution Reconstruction</i>	: Reconstitution par multirésolution
<i>PF</i>	: <i>Peak Frequency</i>	: Fréquence du pic
<i>QMF</i>	: <i>Quadrature Mirror Filter</i>	: Filtres miroirs en quadrature
<i>SNR</i>	: <i>Signal to Noise Ratio</i>	: Rapport signal sur bruit
<i>STFT</i>	: <i>Short Time Fourier Transform</i>	: Transformée de Fourier à court-terme
<i>TDW</i>	: <i>Total Deflexion Width</i>	: Durée de la déflexion totale

INTRODUCTION

Les maladies respiratoires touchent plus de 3 millions de Canadiens de tous les âges ^[1]. Ces maladies incluent l'asthme, les maladies pulmonaires obstructives chroniques (MPOC), le cancer du poumon, la grippe, la pneumonie, etc.

Ces maladies ont des répercussions économiques dramatiques sur le système canadien de soins de santé. En effet, les coûts directs et indirects ont totalisé près de 12,18 milliards pendant l'année 1993 ^[1].

Les bruits respiratoires contiennent des informations significatives permettant d'étudier l'évolution des pathologies du poumon et des voies aériennes ^[2]. L'analyse de ce type de signaux, dans le but de réaliser un système d'aide au diagnostic, a suscité l'intérêt de nombreuses équipes de recherche ^[3-13]. Plusieurs approches ont été proposées pour l'analyse des sons respiratoires, mais il reste encore beaucoup de chemin à faire. En effet, la complexité de ces types de signaux et le manque de standardisation entre les équipes de recherche rendent la tâche plus difficile.

Pour répondre à cette problématique, un projet de recherche a été lancé au département de mathématiques, d'informatique et de génie (DIMG) de l'Université du Québec à Rimouski (UQAR) qui vise à mettre en place un système fiable d'analyse des sons respiratoires. La première phase du projet a été consacrée aux sons adventices continus, dans laquelle de nouvelles méthodes de caractérisation et de modélisation ont été proposées pour la détection et la classification des sibilants ^[14-17].

Notre travail de recherche entre dans cette perspective et vise à construire un système de classifications des sons adventices. Le premier chapitre présente une introduction générale des signaux respiratoires, ainsi que la problématique de cette recherche. Dans le second chapitre, nous présentons les techniques de séparation des sons respiratoires en deux classes, continue et discontinue. Le troisième chapitre introduit l'identification des sons adventices discontinus qui permettra, par la suite, la détection et la quantification des différentes pathologies. De plus, quelques méthodes d'extraction des sons adventices conti-

nus sont présentées dans le quatrième chapitre. Le cinquième chapitre met en évidence les performances des méthodes discutées et fournit les résultats obtenus. Enfin, un bilan et les perspectives de ce travail sont présentés dans la conclusion.

CHAPITRE I

DÉFINITIONS ET CLASSIFICATION DES SONS RESPIRATOIRES

Dans ce chapitre, les définitions et les caractéristiques des différents bruits respiratoires sont introduites. Par la suite, la problématique de cette recherche est décrite. Le système d'analyse des bruits adventices est brièvement présenté par le biais de son schéma.

1.1 Nomenclature et classification des sons respiratoires

L'auscultation est une opération clinique qui consiste à écouter, à l'aide d'un stéthoscope ou simplement à l'oreille, divers bruits produits par les organes du corps humain, tel les poumons et le cœur. Cette méthode de diagnostic médical a été mise au point par René Laennec en 1816 ^[18]. L'auscultation fournit un moyen simple d'observer les sons respiratoires et elle est encore utilisée comme examen complémentaire pour le diagnostic. Cependant, un inconvénient de cette méthode est le caractère subjectif de l'information obtenue ^[19].

Considérant les progrès réalisés dans le domaine de l'enregistrement et de l'analyse des sons, il est désormais possible d'avoir une meilleure connaissance des caractéristiques des sons respiratoires aujourd'hui ^[19]. De nombreux travaux ont été consacrés au mécanisme de production des bruits respiratoires ainsi que leurs relations avec les pathologies ^[2, 20, 21]. Pour réaliser une classification objective des sons respiratoires, il faut définir des critères acoustiques fiables et robustes.

Les sons respiratoires incluent tous les sons perceptibles lors de la respiration. Ils incluent donc, les sons entendus au niveau de la bouche, la trachée ou la cage thoracique.

Ces sons se divisent en deux grandes classes : les sons normaux et les sons adventices. Les sons adventices se superposent aux sons normaux et sont, le plus souvent, des indicateurs de maladies cardio-respiratoires ^[20]. L'analyse des sons adventices permettrait d'étudier l'évolution des maladies bronchiques ou pulmonaires ^[2].

Les sons adventices sont subdivisés en deux classes : continus ou discontinus, c'est-à-dire d'une durée supérieure ou inférieure à 250 ms ^[18]. La classe des sons continus contient les "sibilants" et les "ronchus". Cependant, la classe des sons discontinus contient les "crépitants fins" et les "gros crépitants" ^[2,22]. La figure 1.1 représente le diagramme de classification des sons respiratoires.

1.1.1 Sons normaux

Les sons normaux pourraient être distingués par l'emplacement de la prise de son (voir figure 1.2). Le bruit trachéal est entendu au niveau de la trachée et est produit par la turbulence de l'air lors de la respiration normale. C'est un son intense possédant une haute tonalité et une pause distincte entre deux phases respiratoires. La différence entre l'inspiration et l'expiration n'est pas évidente.

Le murmure vésiculaire est entendu à la périphérie du poumon et est essentiellement inspiratoire. Dans ce cas, le poumon fonctionne comme un filtre passe-bas qui enlève les composants en fréquence haute. Il s'agit alors d'un son doux à basse tonalité. La limite entre l'inspiration et l'expiration n'est pas claire. Le mécanisme de sa production est plus compliqué et n'est pas connu avec exactitude à l'heure actuelle.

Le bruit trachéo-bronchique et le bruit broncho-vésiculaire ressemblent beaucoup au bruit trachéal et au murmure vésiculaire. Ils sont entendus dans les zones à proximité des grosses voies aériennes (voir la figure 1.2).

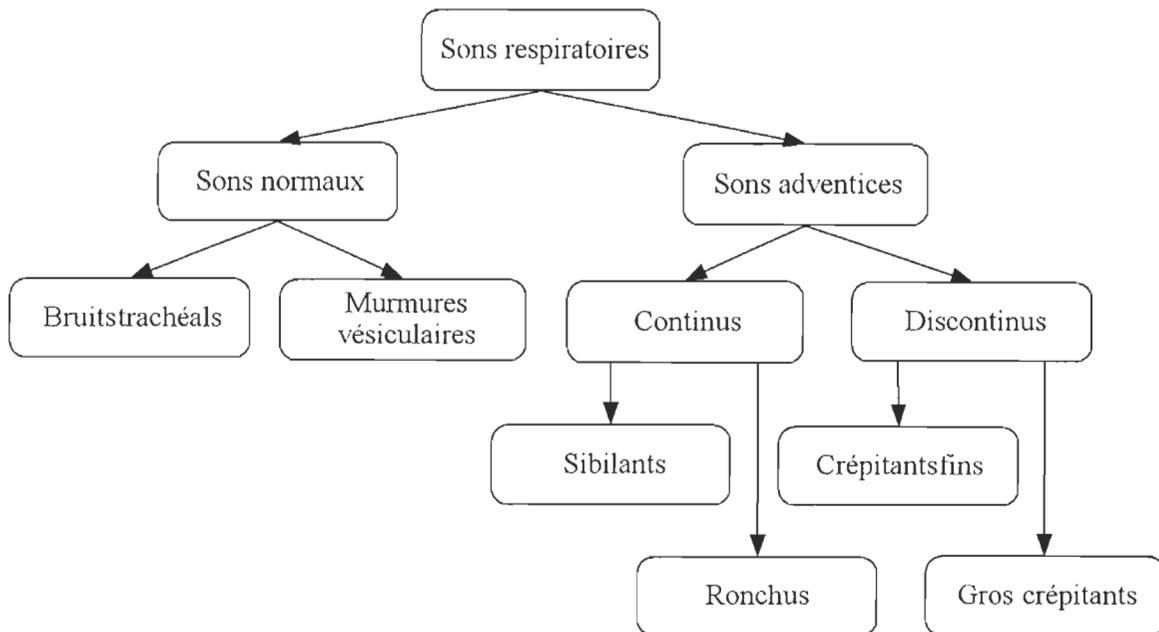


Figure 1.1: Diagramme de classification des sons respiratoires.

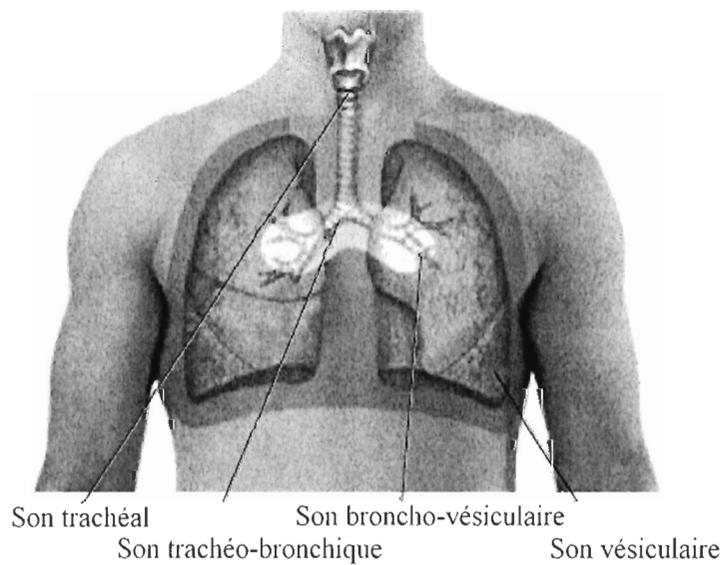


Figure 1.2: Différents types de sons respiratoires normaux. Ils se distinguent par l'emplacement de prise de son ^[21].

Tableau 1.1: Définitions et appellations des sons adventices selon l'American Thoracic Society (Tableau provenant de [23, 24]).

	G.B. et U.S.A.	ALLEMAGNE	FRANCE	Proposition de modification	Forme d'onde en temps étendu	
Discontinuous (< 250 ms)						
	- <u>Fine</u> (high pitched, low amplitude, short duration, IDW* = 0.92 ms, 2CD† = 6.02 ms)	Fine crackles	Feines Rasseln	Râles crépitants	Crépitements fins	
	- <u>Coarse</u> (low pitched, high amplitude, long duration, IDW = 1.25 ms, 2CD = 9.32 ms)	Coarse crackles	Grobes Rasseln	Râles bulleux ou sous-crépitants	Gros crépitements	
Continuous (> 250 ms)						
	- <u>High pitched</u> (dominant frequency > 400 Hz)	Wheezes	Pfeifen	Râles sibilants	Sifflements	
	- <u>Low pitched</u> (dominant frequency < = 200 Hz)	Ronchi	Brummen	Râles ronflants	Ronchus	

* Initial Deflection Width (largeur de déflexion initiale)

† Two Cycle Duration (durée des 2 premiers cycles)

1.1.2 Sons adventices continus

Les sons adventices continus regroupent les sibilants et les ronchus. Il s'agit de bruits anormaux d'une durée supérieure à 250 ms. Ils proviennent de la vibration des parois des voies aériennes étroites (resserrées). Les sibilants sont souvent rencontrés chez les patients ayant une obstruction des voies aériennes, comme pour l'asthme [2].

Selon la normalisation de la terminologie des sons adventices proposée par l'ILSA (*International Lung Sound Association*) et adoptée par l'ATS (*American Thoracic Society*), la fréquence dominante du sibilant est supérieure à 400 Hz et celle du ronchus inférieure à 200 Hz. Le tableau 1.1 montre aussi que leur durée minimale est de 250 ms. Habituellement, les sibilants et les ronchus sont présents dans la phase expiratoire. Par contre, de récentes recherches ont démontré que les descriptions de ces sons adventices ne sont pas assez précises. Les sibilants sont aussi présents dans la phase d'inspiration. D'après CORSA (*Computerized Respiratory Sound Analysis*), la durée acceptable pour les sons adventices discontinus est de 80 à 100 ms. La fréquence dominante d'un sibilant est supérieure à 100 Hz et celle du ronchus est inférieure à 300 Hz [22]. Dans ce projet de recherche, la normalisation proposée par CORSA sera utilisée.

Les sibilants qui contiennent une seule fréquence s'appellent "monophoniques" et ceux qui contiennent plusieurs fréquences s'appellent "polyphoniques". Les sibilants peuvent être écoutés au niveau de la bouche, de la trachée et du sternum. Mais à cause de l'effet atténuateur de la cage thoracique (filtre passe-bas), les sibilants de fréquences élevées pourraient être éliminés [25,26]. Par conséquent, pour l'analyse des sibilants, les sons respiratoires sont souvent enregistrés au niveau de la trachée.

La figure 1.3 représente deux exemples de sons adventices (sibilant et ronchus) ainsi que leurs spectrogrammes respectifs. La fréquence dominante du sibilant (figure 1.3-a) est supérieure à celle du ronchus (figure 1.3-b).

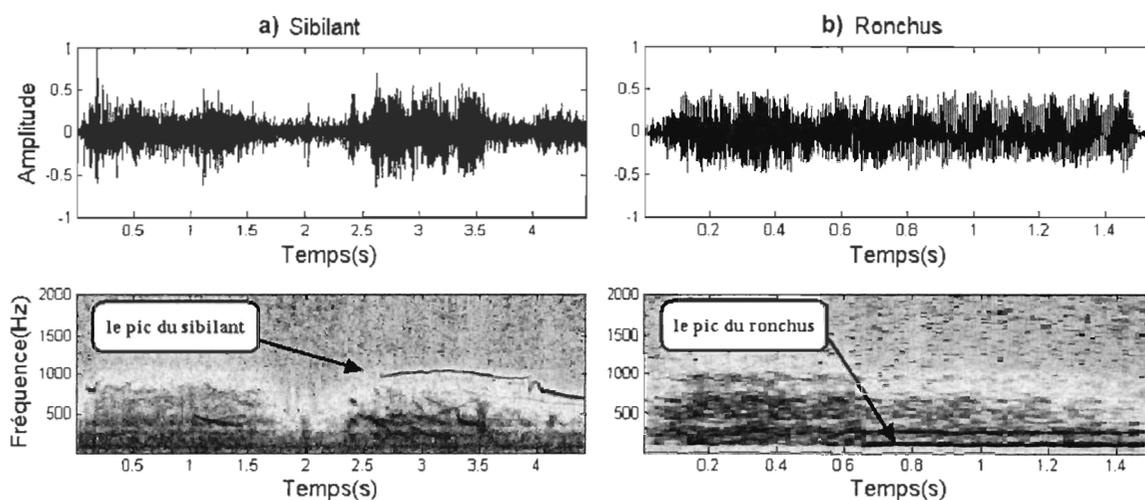


Figure 1.3: Représentation dans le domaine temps et sous forme de spectrogramme des sons adventices continus a) son respiratoire avec sibilant, b) son respiratoire avec ronchus. Il est à noter que la fréquence dominante du ronchus est inférieure à celle du sibilant.

1.1.3 Sons adventices discontinus

Les crépitants sont des sons adventices discontinus dont la durée ne dépasse pas 250 ms. De ce fait, ils sont qualifiés de transitoires ou discontinus. Les crépitants sont, généralement, divisés en deux sous-classes : les "crépitants fins" et les "gros crépitants" (voir le tableau 1.1). La figure 1.4 représente deux exemples de sons adventices (crépitant fin et gros crépitant) et leurs spectrogrammes respectifs.

Afin de distinguer les crépitants fins des gros crépitants, on utilise le plus souvent, la largeur de la déflexion initiale IDW (*Initial Deflexion Width*), et la durée des deux premiers cycles 2CD (*Two Cycle Duration*) (voir la figure 1.5). Selon la définition de l'ATS, les durées moyennes de IDW et 2CD des crépitants fins sont respectivement de 0.7 ms et 5 ms, et celles des gros crépitants sont respectivement de 1.5 ms et 10 ms. C'est-à-dire, les crépitants fins sont caractérisés par une fréquence plus importante occupant une courte durée, alors que les gros crépitants sont caractérisés par une fréquence plus faible occupant une durée plus longue ^[20]. CORSA propose une autre définition selon laquelle la 2CD du crépitant fin est inférieure à 10 ms et celle du gros crépitant est supérieure à

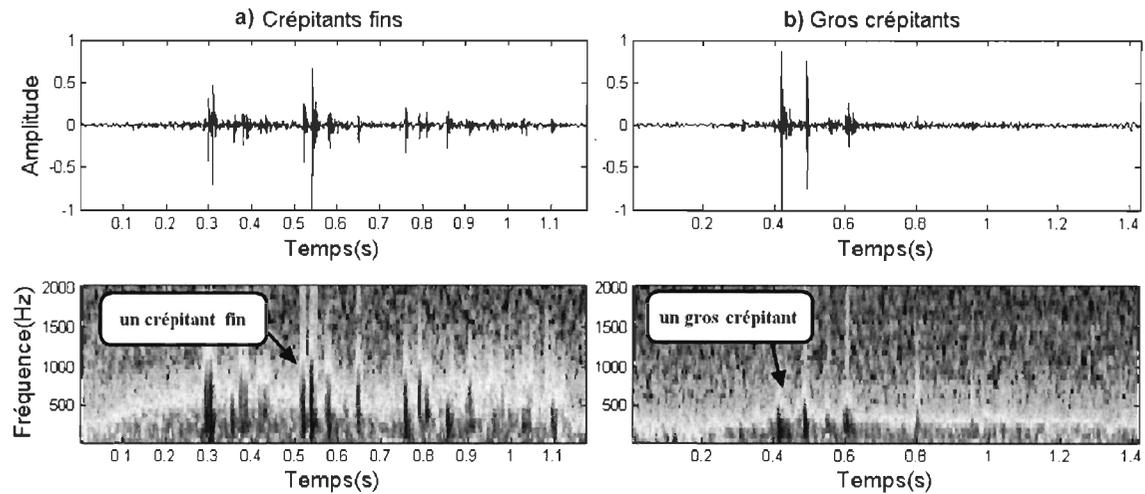


Figure 1.4: Représentation dans le domaine temps et sous forme de spectrogramme des sons adventices discontinus a) son respiratoire avec crépitants fins, b) son respiratoire avec gros crépitants.

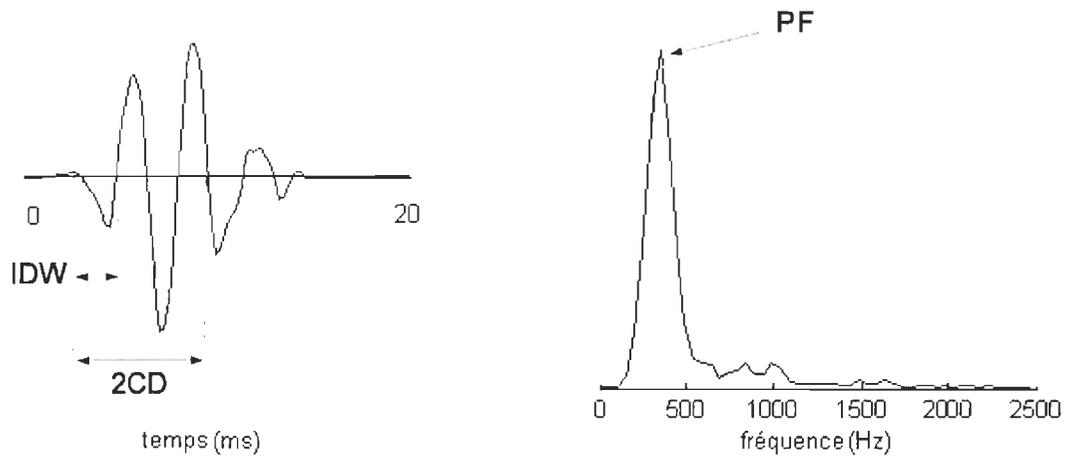


Figure 1.5: Définitions de la largeur de la déflexion initiale (IDW), la durée des deux premiers cycles (2CD) et la fréquence du pic (PF) pour les crépitants.

Tableau 1.2: Caractéristiques des sons adventices et les pathologies leurs sont associés [20,22,28].

Nomenclature	Caractéristique	Description temporelle	Description fréquentielle	Pathologies principales
Sibilant	Continu	> 100 ms	> 100 Hz	Asthme
Ronchus	Continu	> 100 ms	< 300 Hz	Bronchite chronique
Crépitant fin	Discontinu	2CD < 10 ms	100-2000 Hz	Fibrose interstitielle, Pneumonie
Gros crépitant	Discontinu	2CD > 10 ms	100-2000 Hz	Broncho-pneumonie, Bronchiectasies

10 ms [22,27].

L'analyse des crépitants donne des informations très utiles pour le diagnostic des maladies pulmonaires. Le type de crépitant indique le nature de la maladie, alors que le nombre de crépitants dans un cycle respiratoire reflète sa sévérité. En effet, les gros crépitants sont signe d'une bronchopneumonie ou de bronchiectasies ; tandis que, les crépitants fins sont signe d'une fibrose interstitielle ou d'une pneumonie [28].

Les crépitants peuvent être classés selon leur position dans le cycle de la respiration. Les crépitants fins sont des événements qui apparaissent exclusivement dans la phase d'inspiration et sont souvent présents dans la mi-inspiration et à la fin de l'inspiration. Les gros crépitants peuvent être trouvés au début de la phase d'inspiration, et parfois dans la phase d'expiration [29]. Le tableau 1.2 résume les caractéristiques des sons adventices et les pathologies qui leurs sont associées.

1.2 Base de données utilisée

Les sons respiratoires proviennent des bases de données ASTRA [19] et RALE [30], ainsi que des CD-ROM de deux livres traitant des sons respiratoires [31,32]. Les sons utilisés sont déjà répertoriés en bruits trachéaux, bruits vésiculaires, sibilants, crépitants fins et gros crépitants. Ces sons audio sont enregistrés dans différentes conditions. En effet, nous ignorons toutes les informations reliées aux conditions d'enregistrement (type de capteur, filtre

anti-repliement, ...) et aux patients (âge, sexe, taille, ...). Ce manque d'informations ne constitue pas vraiment un inconvénient car le but est de réaliser un système robuste, fiable et insensible aux conditions de mesure. En vue d'harmoniser le traitement de ses fichiers, nous réalisons trois opérations de pré-traitement : 1) la normalisation des amplitudes, 2) le re-échantillonnage pour les ramener à la même fréquence d'échantillonnage, et 3) la segmentation manuellement pour avoir un seul cycle respiratoire par fichier.

1.3 Problématique de recherche

Les sons respiratoires adventices sont synonymes du dysfonctionnement du système respiratoire et sont associés aux différentes maladies pulmonaires. La problématique de recherche vise à développer des techniques de reconnaissance automatique des sons adventices dans le but de réaliser un système d'aide au diagnostic des maladies. Ce type de systèmes a suscité l'intérêt de plusieurs équipes de recherche mais ils sont toujours en phase de développement. La complexité des signaux respiratoires et le manque de standardisation entre les équipes de recherche freinent considérablement la recherche dans ce domaine.

Les figures 1.3 et 1.4 montrent que les caractéristiques temporelles et fréquentielles des deux classes de sons adventices sont très différentes. En effet, les sons adventices continus sont de nature stationnaire ; ils sont bien localisés en fréquence mais couvrent un intervalle de temps relativement long. Par contre, les sons adventices discontinus sont de nature non-stationnaire ; ils sont bien localisés dans le temps mais très étalés en fréquence. Il devient difficile de développer une technique permettant la reconnaissance simultanée des différents sons adventices. Dans le cas de l'analyse spectrale, il faut analyser les sons adventices sur des fenêtres temporelles de largeurs extrêmement différentes, dépendamment de leur nature stationnaire ou non-stationnaire.

1.4 Objectifs

Ce projet de recherche a pour but de mettre en place un système de reconnaissance automatique des sons adventices. Il consiste à améliorer les méthodes existantes d'analyse des sons respiratoires et de proposer de nouvelles solutions. Nous nous intéressons particulièrement aux sons discontinus, à savoir les crépitants fins et les gros crépitants. Nous examinons aussi les performances de nouvelles techniques d'extraction des sibilants.

1.4.1 Analyse des crépitants

Les crépitants sont des signaux explosifs, transitoires et surajoutés aux bruits normaux. Ils sont indicateurs de dysfonctionnement du système ventilatoire. Les crépitants sont caractérisés par l'instant de leur apparition dans le cycle respiratoire (au début de l'inspiration ou à la fin de l'expiration etc.), leur nombre (élevé ou faible), et leur tonalité (basse ou haute) [33]. En raison de leur faible puissance et de leur courte durée, il est très difficile de les analyser en présence de sons normaux. Dans le présent mémoire, l'analyse des crépitants se divise en quatre étapes, à savoir : 1) Séparation : extraire les crépitants du son respiratoire normal, 2) Débruitage : filtrer le signal contenant les crépitants pour réduire le bruit, 3) Détection : localiser la position du crépitant dans le temps, et 4) Classification : catégoriser les crépitants en deux classes, gros ou fin. La figure 1.6 montre le schéma du système d'analyse des crépitants.

- **Séparation stationnaire/non-stationnaire** : Lors de cette étape, les signaux sont divisés en deux classes : continue et discontinue. Plusieurs techniques seront évaluées : filtre non-linéaire (ST-NST) [35-37], filtre à base de la transformée par ondelettes (WTST-NST) [38], filtre utilisant le système d'inférences floue (GFST-NST) [39] et un filtre basé sur les paquets d'ondelettes (WPST-NST) que nous avons proposé [40].
- **Débruitage** : Vise à réduire le niveau de bruit de la composante non-stationnaire du signal respiratoire.

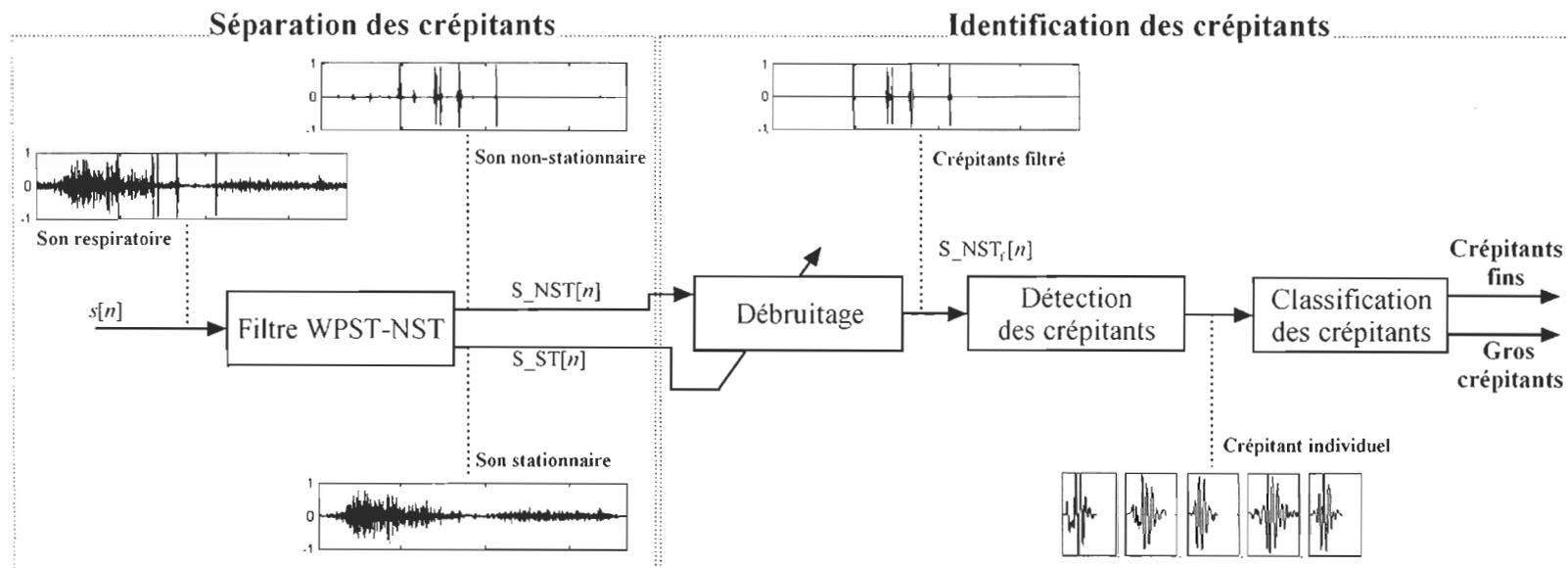


Figure 1.6: Schéma du système d'analyse des crépitants [34].

- **Détection** : Cette étape consiste à détecter la présence d'un crépitant en déterminant l'instant de son apparition. L'algorithme de détection est basé sur le paramètre "Dimension Fractale" (FD) [41]. Un nouvel algorithme d'énumération des crépitants est proposé dans ce mémoire [34].
- **Classification** : Cette étape consiste à extraire les paramètres caractéristiques du crépitant afin de les différencier. Différentes combinaisons de paramètres sont testées afin de déterminer la meilleure.

1.4.2 Analyse des sibilants

Pour compléter les travaux antérieurs de l'équipe [14-17], nous allons tester et évaluer quatre autres méthodes de détection des sibilants basées globalement sur les spectrogrammes : 1) Laplacien d'une gaussienne [42], 2) modélisation auditive [43], 3) détection d'épisodes de sibilants [44], et 4) représentation par "Chirplets" [45].

CHAPITRE II

EXTRACTION DES CRÉPITANTS DES SONS VÉSICULAIRES

Ce chapitre concerne l'étape de séparation des sons continus et discontinus, qualifiés aussi de stationnaires et non-stationnaires. Dans cette étape, les crépitants seront extraits des sons vésiculaires afin de les reconnaître ultérieurement. Nous présentons trois techniques de séparation rencontrées dans la littérature ; elles sont respectivement basées sur : le filtre adaptatif, l'inférences floue et la transformée par ondelettes. Enfin, nous proposerons une nouvelle technique basée sur les paquets d'ondelettes.

2.1 Principe de séparation stationnaire et non-stationnaire des sons respiratoires

Les crépitants sont des sons respiratoires explosifs de courte durée, qui sont associés le plus souvent à un dysfonctionnement pulmonaire. Il s'agit de sons transitoires surajoutés aux bruits respiratoires normaux. Le processus de séparation des crépitants des sons vésiculaires constitue une étape importante dans l'analyse des sons respiratoires. C'est une tâche laborieuse en raison de la faible énergie et de la durée relativement courte des crépitants.

Le principe de séparation des sons respiratoires est basé sur les hypothèses suivantes : les sons adventices discontinus (crépitants) peuvent être considérés comme des sons non-stationnaires alors que les sons normaux (vésiculaires) et les sons adventices continus (asthme et rhonchus) sont considérés comme stationnaires.

Dans la section qui suit, nous décrivons en détails trois techniques de séparation basées

respectivement sur : le filtrage adaptatif (ST-NST) [35,37], le système d'inférence floue (GFST-NST) [39], et la transformée en ondelettes (WTST-NST) [38]. Nous décrivons aussi le nouveau filtre, basé sur les paquets d'ondelettes (WPST-NST), que nous avons proposé récemment [40]. Le filtre proposé offre de meilleures performances comparativement aux méthodes existantes.

2.2 Extraction des crépitants par filtre non-linéaire

Le filtre ST-NST se base sur une approche non-linéaire pour séparer les signaux non-stationnaires des signaux stationnaires [35]. On suppose que le signal d'entrée $x[n]$ est la somme de deux signaux : un signal stationnaire qui peut être exprimé par le modèle prédictif et un signal non-stationnaire composé d'ondes impulsives aléatoires de faible taux d'apparition. En premier lieu, une introduction du filtre prédictif est présentée.

2.2.1 Filtre prédictif

Le filtre prédictif consiste à prédire le prochain échantillon d'un processus stochastique stationnaire à partir d'un nombre restreint de ses échantillons précédents.

$$\tilde{x}[n] = \sum_{k=1}^M a_k x[n-k] \quad n = 1, 2, \dots, N \quad (2.1)$$

où $\tilde{x}[n]$ est la prédiction du signal $x[n]$ et a_k représente les coefficients du filtre prédictif d'ordre M . La figure 2.1 montre un exemple de filtre predictif basé sur un filtre adaptatif. L'erreur de prédiction $e_p[n]$ pour le $n^{\text{ème}}$ échantillon $x[n]$ est donnée par la différence entre l'échantillon actuel et sa valeur prédite :

$$e_p[n] = x[n] - \tilde{x}[n] = x[n] - \sum_{k=1}^M a_k x[n-k] \quad (2.2)$$

Les coefficients a_k du filtre sont ajustés par l'algorithme LMS (*Least Mean Square*). C'est une technique itérative basée sur la minimisation de l'erreur quadratique moyenne. En

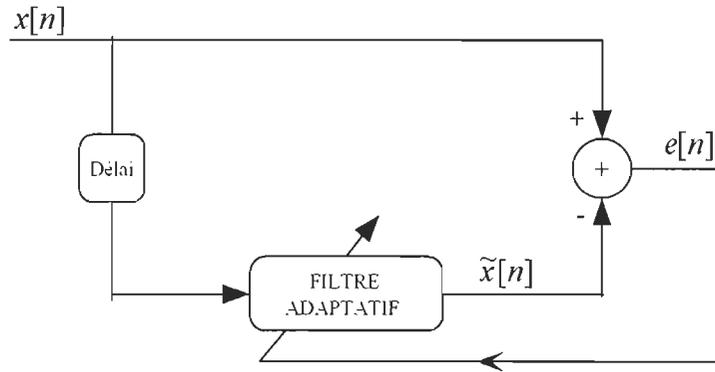


Figure 2.1: Structure du filtre prédictif ^[46].

fait, les coefficients du filtre sont ajustés à chaque itération ^[46].

$$a_k[n + 1] = a_k[n] + 2\mu e[n]x[n - k] \quad k = 0, 1, \dots, M - 1 \quad (2.3)$$

où μ est une constante positive qui détermine la vitesse de convergence et la stabilité du filtre. Une valeur faible de μ assure la stabilité mais un temps de convergence trop important et vice et versa.

2.2.2 Filtre de séparation non-linéaire ST-NST

Ce filtre de séparation stationnaire et non-stationnaire (ST-NST) a été proposé par Ono et al. ^[35] pour extraire les crépitants des sons vésiculaires. La figure 2.2 représente le diagramme de ce filtre qui est principalement composé d'un filtre prédictif et d'une fonction non-linéaire $F(\cdot)$.

Dans la section précédente, nous avons vu que le filtre prédictif permet d'obtenir l'estimation de l'échantillon d'un signal stationnaire en utilisant les M derniers échantillons.

$$\tilde{y}[n] = \sum_{k=1}^M a_k[n]\hat{y}[n - k] \quad (2.4)$$

où a_k et M représentent respectivement les coefficients et l'ordre du filtre prédictif.

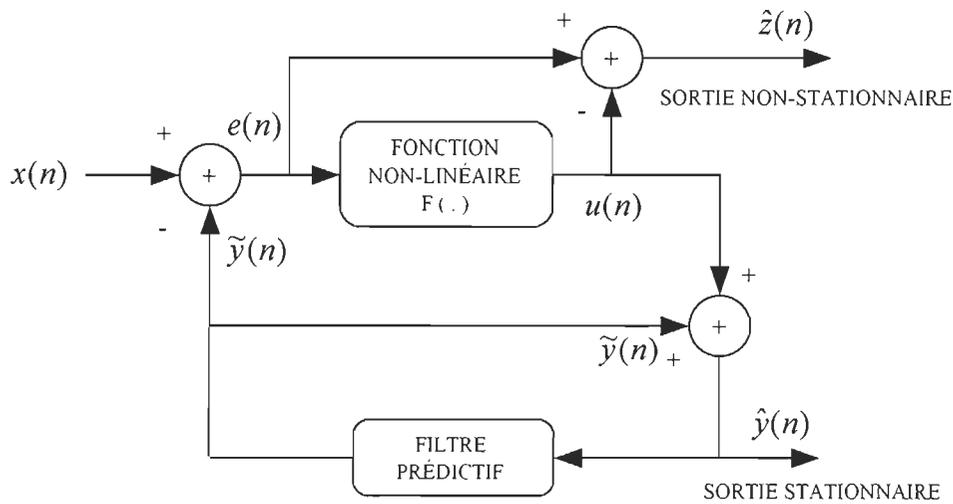


Figure 2.2: Schéma du filtre ST-NST ^[35] utilisant le filtre prédictif de la figure 2.1.

En se référant au diagramme de la figure 2.2, on constate que l'erreur $e[n]$ est définie par la différence entre l'entrée $x[n]$ et la prédiction de la composante stationnaire $\tilde{y}[n]$.

$$e[n] = x[n] - \tilde{y}[n] \quad (2.5)$$

Ensuite, l'erreur est traitée par la fonction non-linéaire $F(\cdot)$, définie par le tracé de la figure 2.3, dans le but de déterminer une sortie $u[n]$ faisant partie de la composante stationnaire du signal.

$$u[n] = F(e[n]) \quad (2.6)$$

Ainsi, la sortie stationnaire $\hat{y}[n]$ et la sortie non-stationnaire $\hat{z}[n]$ sont définies respectivement par les équations 2.7 et 2.8 :

$$\hat{y}[n] = \tilde{y}[n] + u[n] \quad (2.7)$$

$$\hat{z}[n] = e[n] - u[n] \quad (2.8)$$

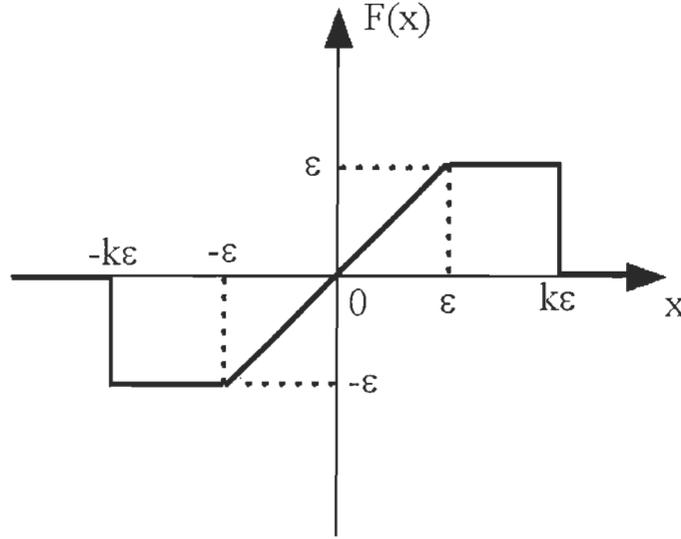


Figure 2.3: Fonction non-linéaire utilisée par le filtre ST-NST ^[35].

En se basant sur la structure du filtre (figure 2.2) et la description de la fonction non-linéaire (figure 2.3), il est facile de vérifier que :

- Si $|e[n]| < \varepsilon$, alors la sortie stationnaire est égale au signal d'entrée, tandis que la sortie non-stationnaire est nulle : $\hat{y}[n] = x[n]$ et $\hat{z}[n] = 0$;
- Si $|e[n]| > k\varepsilon$, alors la sortie stationnaire est égale à la sortie du filtre adaptatif, tandis que la sortie non-stationnaire est égale à l'erreur prédictive : $\hat{y}[n] = \tilde{y}[n]$ et $\hat{z}[n] = e[n]$;
- Si $\varepsilon < |e[n]| < k\varepsilon$, alors le signal d'entrée est traité comme un signal semi-stationnaire : $\hat{y}[n] = \tilde{y}[n] + \varepsilon$ et $\hat{z}[n] = e[n] - \varepsilon$.

où le paramètre ε est déterminé de sorte que la probabilité de détection du signal non-stationnaire est définie par une valeur fixe γ . La fonction $p(x)$ représente la densité de probabilité du signal d'erreur.

$$\int_{-\varepsilon}^{\varepsilon} p(x) dx = 1 - \gamma \quad (2.9)$$

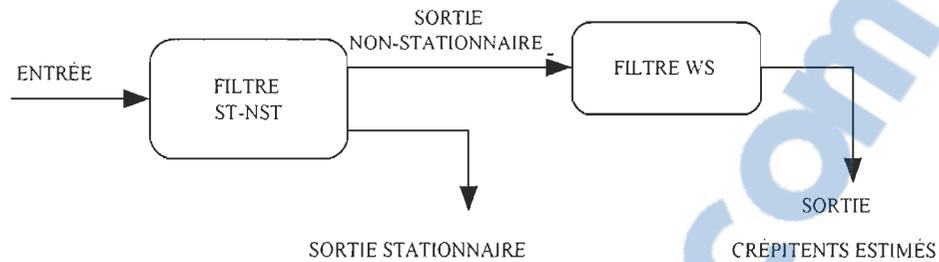


Figure 2.4: Diagramme d'un filtre numérique pour l'extraction des crépitants [36].

Une version modifiée de ce filtre a été proposée deux ans plus tard par la même équipe [36]. Ils ont ajouté un filtre de post-traitement à la sortie non-stationnaire (figure 2.4), capable de réduire les bruits selon l'intervalle de temps séparant deux passages par zéro successifs des signaux non-stationnaires. Une autre version du filtre ST-NST, utilisant les statistiques d'ordre supérieur comme seuil de discrimination, a été proposée plus tard [37]. Il s'est avéré qu'elle est plus robuste pour réduire les bruits blancs. Toutefois, dans les trois modèles, nous avons besoin de la définition empirique de l'ensemble des coefficients.

2.3 Extraction des crépitants par logique floue

Un filtre GFST-NST (*Generalized Fuzzy rule-based STationary-NonSTationary filter*) basée sur la logique floue a été proposé par Tolia et al. [39] pour extraire les crépitants des sons vésiculaires. En fait, cet outil est composé de deux systèmes d'inférence floue à base d'un réseau adaptatif ANFIS (*Adaptive-Network-based Fuzzy Inference System*). Avant de décrire le fonctionnement du filtre, nous allons introduire les notions de base de la logique floue.

2.3.1 Concept de base de la logique floue

Le plus souvent, une personne n'a pas besoin d'information précise ou numérique pour prendre une décision. Par exemple, est-ce que la température de l'eau est adéquate

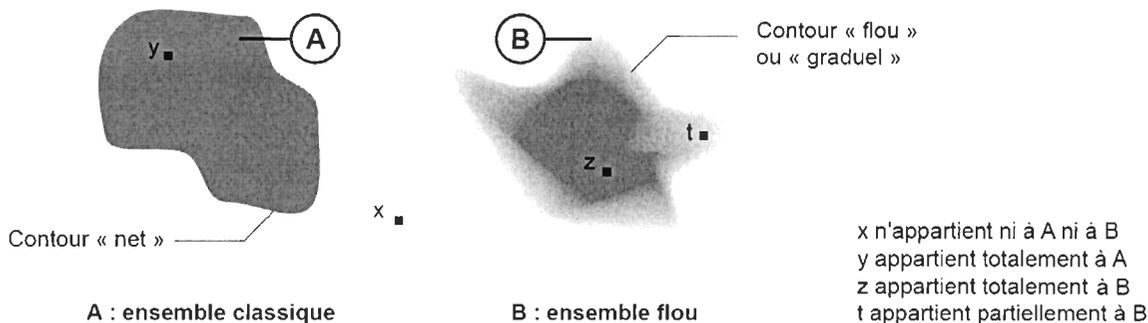


Figure 2.5: Comparaison d'un ensemble classique et d'un ensemble flou ^[47]. Les contours d'un ensemble classique sont "nets" alors que ceux de l'ensemble flou sont "flous" ou "graduels".

pour prendre une douche? Est-ce que ce steak est suffisamment cuit pour être mangé? Pour répondre à ce type de questions que le concept de logique floue a été proposé par le professeur Lotfi Zadeh. Contrairement à la logique bivalente, la logique floue permet de manipuler des valeurs de vérité autres que le "vrai" et le "faux" absolus. Elle sert donc à représenter des connaissances incertaines et imprécises, c'est à dire que la vérité est exprimée par une manière de degré. Aujourd'hui, l'approche par logique floue est universellement utilisée dans les systèmes de contrôle ayant des entrées imprécises ou contaminées. Elle permet de résoudre des problèmes de façon simple et "humanisée". Le concept de base de la logique floue comprend quatre parties : l'ensemble flou, la fonction d'appartenance, l'opérateur flou et les règles d'inférence.

a) Ensemble flou

Dans la théorie des ensembles classiques, un élément appartient ou n'appartient pas à un ensemble. Parfois, cette notion ne permet cependant pas de répondre à des situations simples mais rencontrées fréquemment. En logique floue, la théorie des ensembles flous repose sur la notion d'appartenance partielle, les contours de chaque ensemble flou sont "flous" ou "graduels" ^[47] (voir figure 2.5).

Prenons comme exemple, la définition de la température de l'eau. En logique clas-

sique, l'eau est froide ou chaude. Nous conviendrons par exemple que l'eau est chaude si sa température est supérieure à 40°C , puis l'eau est froide si sa température est inférieure à 40°C . Alors la température de l'eau est divisée en deux intervalles $[0^{\circ}\text{C}, 40^{\circ}\text{C}[$ et $[40^{\circ}\text{C}, 100^{\circ}\text{C}]$. Dans ce cas, 40°C est le point de délimitation entre les deux ensembles, l'eau de 39°C est froide mais celle de 40°C est chaude. La discontinuité de cette définition n'est pas naturelle (voir la figure 2.6-a).

En logique floue, l'ensemble est défini par un degré d'appartenance, nous pouvons avoir une délimitation plus douce (graduelle). Si la température était de 80°C , l'eau est chaude absolument, par contre si la température était de 40°C , l'eau n'est ni vraiment chaude ni vraiment froide (voir la figure 2.6-b).

b) Fonction d'appartenance

Un ensemble flou est défini par sa "fonction d'appartenance", qui correspond à la notion de "fonction caractéristique" en logique classique ^[47].

Supposons que nous voulions définir la température de l'eau par trois ensembles : "l'eau froide", "l'eau tiède" et "l'eau chaude". En logique classique, nous conviendrons que la température de l'eau tiède est comprise entre 25°C et 55°C . La fonction caractéristique de l'ensemble donne "0" pour les températures hors de l'intervalle $[25^{\circ}\text{C} \ 55^{\circ}\text{C}]$ et "1" pour celles qui sont à l'intérieur. L'ensemble flou de "l'eau tiède" sera défini par une "fonction d'appartenance" qui diffère de la fonction caractéristique par le fait qu'elle peut prendre n'importe quelle valeur dans l'intervalle $[0, 1]$. À chaque température possible correspondra un "degré d'appartenance" à l'ensemble flou de "l'eau tiède", compris entre 0 et 1. Les ensembles flous pourraient être entrelacés dans certains intervalles. Par exemple, si la température de l'eau était de 25°C , le degré d'appartenance à l'ensemble de "l'eau froide" est 0.5 qui est le même pour son appartenance à l'ensemble de "l'eau tiède" (voir figure 2.7). Le tableau 2.1 donne le degré d'appartenance pour les températures 25°C , 40°C et 55°C pour les trois ensembles flous définis par les fonctions d'appartenance de la figure 2.7-b.

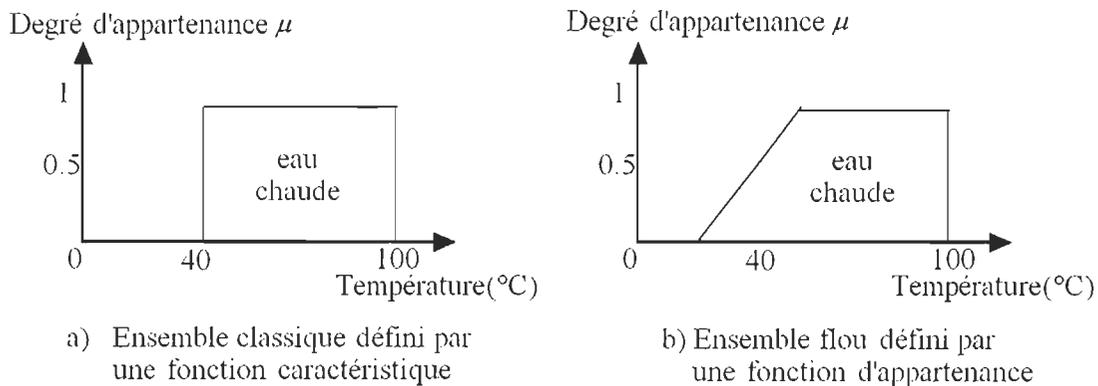


Figure 2.6: Définitions possibles de l'ensemble de "l'eau chaude".

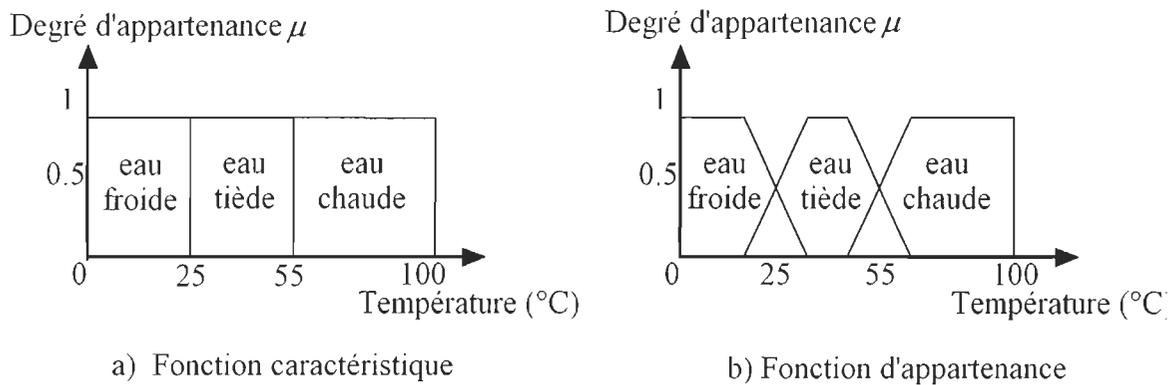


Figure 2.7: Exemples de fonctions caractéristiques et de fonctions d'appartenance.

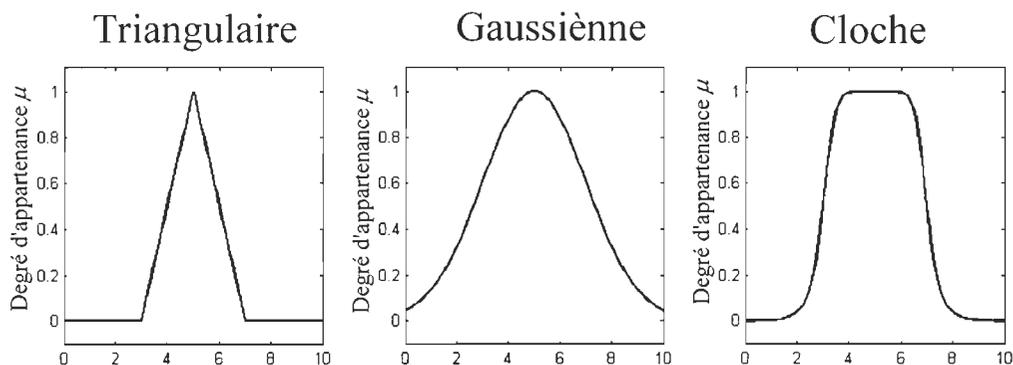


Figure 2.8: Trois exemples de fonctions d'appartenance.

Tableau 2.1: Degré d'appartenance à trois ensembles flous.

Température	Degré d'appartenance		
	Eau froide	Eau tiède	Eau chaude
25°C	0.5	0.5	0
40°C	0	1	0
55°C	0	0.5	0.5

Cet exemple montre la gradualité introduite par la logique floue. L'eau ayant une température de 55°C appartient simultanément à l'ensemble de "l'eau tiède" avec un degré de 0.5 et à l'ensemble de "l'eau chaude" avec un degré de 0.5 quand même. Le terme *température* s'appelle *variable linguistique*, alors que les termes *chaude* ou *froide*, définis par les fonctions d'appartenance, s'appellent *valeurs linguistiques*.

Les fonctions d'appartenance peuvent théoriquement prendre n'importe quelle forme. Toutefois, elles sont souvent définies par la fonction triangulaire, gaussienne ou cloche à cause de leur simplicité (voir la figure 2.8).

c) Opérateurs de logique floue

La logique booléenne est un cas particulier de la logique floue. En logique classique, il y a deux états logiques : faux et vrai. En logique floue, la variable floue représente le degré de la vérité par une valeur comprise entre 0 et 1. Les opérateurs "Complément" (**NON**), "Union" (**OU**) et "Intersection" (**ET**) sont définies par les équations suivantes :

$$\begin{aligned}
 \mu_{\bar{A}}(x) &= 1 - \mu_A(x) && \text{NON} \\
 \mu_{A \cup B}(x) &= \max(\mu_A(x), \mu_B(x)) && \text{OU} \\
 \mu_{A \cap B}(x) &= \min(\mu_A(x), \mu_B(x)) && \text{ET}
 \end{aligned}
 \tag{2.10}$$

d) Règles d'inférence

Le système logique floue fonctionne de manière "humanisée", il est souvent accompli par une base de règles d'inférence. Les règles d'inférence peuvent être exprimées sous la forme : SI "prédicat", ALORS "conclusion" ^[47]. Donc, une règle de type :

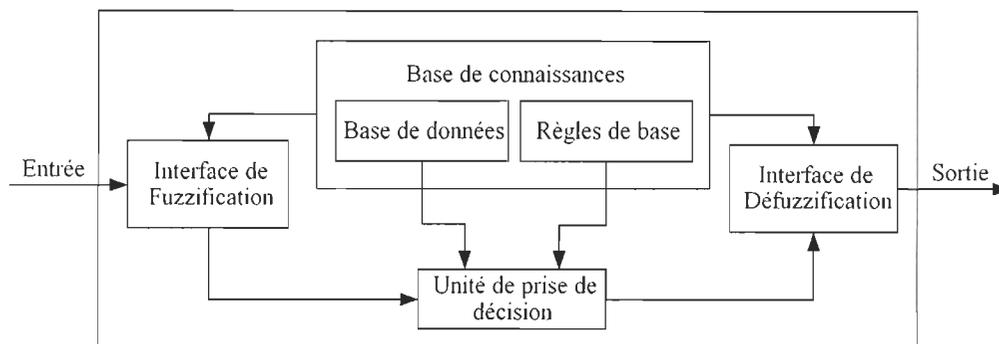


Figure 2.9: Schéma du système d'inférence floue (FIS) ^[48].

SI X est A, ALORS Y est B.

où X et Y sont deux variables linguistiques, A et B sont les étiquettes linguistiques caractérisés par les fonctions d'appartenance. La partie "Si X est A" s'appelle *prédicat*, *antécédent* ou *prémisse*, la partie "Alors Y est B" s'appelle *conséquence* ou *conclusion*. Par exemple, SI la température est élevée dans la chambre, ALORS fermer le chauffage. Ces règles permettent de relier les variables floues d'entrée aux variables floues de sortie à l'aide de différents opérateurs. Normalement, les règles d'inférence sont définies par le concepteur (expert) du système en fonction de son expérience ou de ses besoins.

2.3.2 Système d'inférence floue (FIS)

Un système d'inférence floue FIS (*Fuzzy Inference System*) se compose de cinq parties : 1) une base de règles contenant un certain nombre de règles floues (SI-ALORS) ; 2) une base de données qui définit les fonctions d'appartenance ; 3) une unité de prise de décision qui effectue l'opération d'inférence par les règles floues ; 4) l'interface de fuzzification qui compare les entrées nettes avec les fonctions d'appartenance pour avoir un degré d'appartenance ; 5) l'interface de défuzzification qui transforme les résultats flous en une sortie nette (voir la figure 2.9) ^[48].

La base de règles et la base de données sont souvent regroupées pour former la base de connaissances ^[48] ; la sortie de la fuzzification s'appelle le degré d'appartenance ; les poids

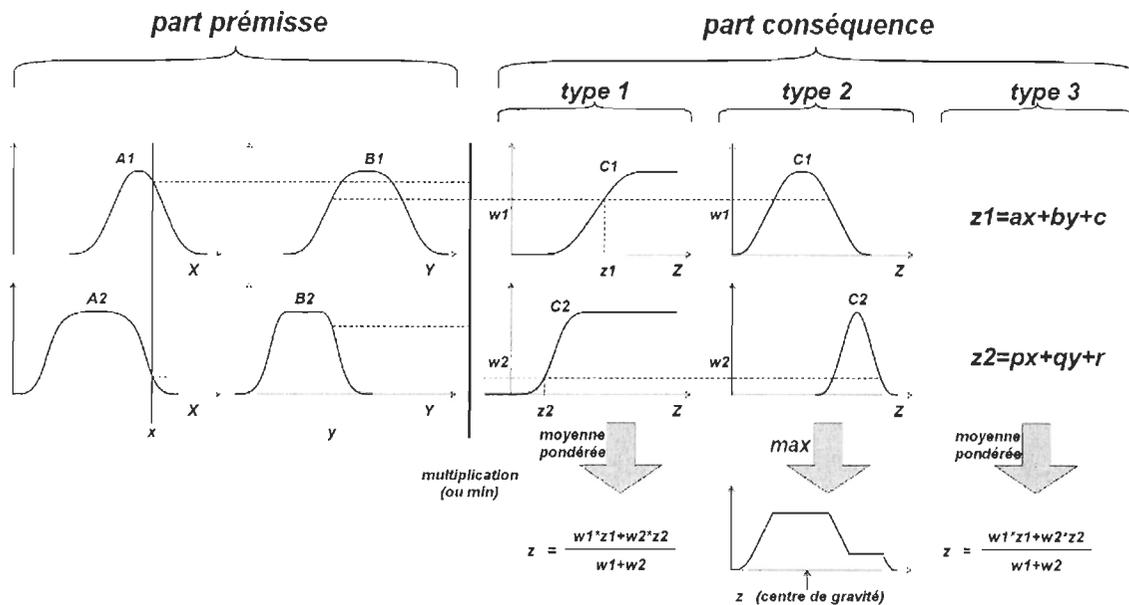


Figure 2.10: Types de règles floues SI-ALORS et les mécanismes de raisonnement flou communément utilisés [48].

des règles sont calculés par un opérateur T-norm, normalement réalisé en déterminant le minimum des degrés d'appartenance.

L'opération d'inférence, appelée aussi le raisonnement flou, dresse une liaison entre l'entrée et la sortie en vertu des règles d'inférence. On rencontre le plus souvent trois types de FIS qui diffèrent par leurs règles floues et leurs mécanismes de raisonnement [48] :

Type 1 : La sortie globale est la moyenne pondérée des sorties conséquentes, induite par les poids des règles et les fonctions d'appartenance. Les fonctions d'appartenance à la sortie devraient être monotone non-décroissant.

Type 2 : La sortie globale est obtenue en appliquant l'opérateur "max" aux sorties floues. Il existe plusieurs méthodes pour calculer la sortie finale, par exemple le centre de gravité ou la moyenne maximale.

Type 3 : La sortie globale est la moyenne pondérée des sorties conséquentes qui sont dérivées par combinaison linéaire des variables d'entrée. Il s'appelle aussi Sugeno-FIS [49].

La figure 2.10 utilise deux entrées et deux règles d'inférence floue pour montrer les différents types de règles floues et des mécanismes de raisonnement. Nous pouvons voir que la différence principale entre les trois méthodes réside dans la façon de défuzzifier.

2.3.3 Système d'inférence floue à base d'un réseau adaptatif (ANFIS)

Le système d'inférence floue à base d'un réseau adaptatif ANFIS (*Adaptive Neuro-Fuzzy Inference System*) est un système adaptatif qui est capable d'ajuster ses fonctions d'appartenance par apprentissage ^[48] afin d'obtenir la sortie désirée. La méthode d'apprentissage utilisée est similaire à celle des réseaux de neurones : les coefficients prémisses (paramètres de la fonction d'appartenance d'entrée) sont ajustés par l'algorithme rétro-propagation ; les coefficients conséquences (paramètres de la fonction d'appartenance de sortie) sont ajustés par l'algorithme LMS (*Least Mean Square*). Il s'agit, donc, d'une méthode d'apprentissage hybride.

Prenons un exemple du système Sugeno-ANFIS ayant deux entrées et deux fonctions d'appartenance pour chaque entrée. Les quatre règles d'inférence sont définies par :

Règle 1 : SI x est A_1 et y est B_1 , ALORS $f_1 = p_1x + q_1y + r_1$.

Règle 2 : SI x est A_1 et y est B_2 , ALORS $f_2 = p_2x + q_2y + r_2$.

Règle 3 : SI x est A_2 et y est B_1 , ALORS $f_3 = p_3x + q_3y + r_3$.

Règle 4 : SI x est A_2 et y est B_2 , ALORS $f_4 = p_4x + q_4y + r_4$.

La figure 2.11 donne la structure de ce réseau ANFIS, qui est un réseau à cinq couches. Les fonctions de chaque couche sont décrites comme suit :

Couche fuzzification : Quantification des entrées x et y aux valeurs floues par les fonctions d'appartenance.

$$O_i^1 = \mu_{A_i}(x, \alpha_i, \beta_i, \gamma_i) = \frac{1}{1 + \left[\left(\frac{x - \gamma_i}{\alpha_i} \right)^2 \right]^{\beta_i}} \quad i = 1, 2; \quad (2.11)$$

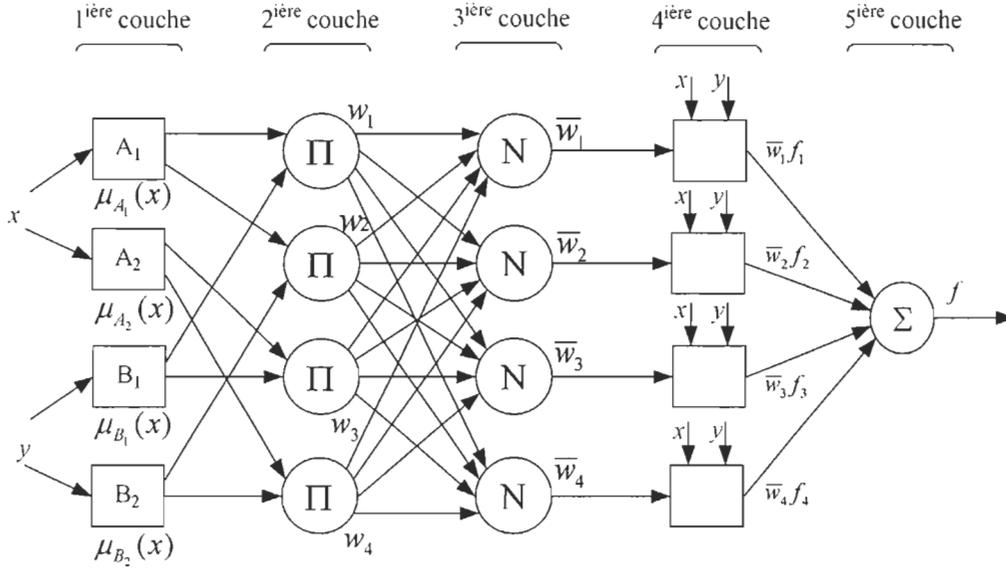


Figure 2.11: Système ANFIS utilisé dans le filtre GFST–NST [39,48].

$$O_i^1 = \mu_{B_{i-2}}(y, \alpha_i, \beta_i, \gamma_i) = \frac{1}{1 + \left[\left(\frac{y - \gamma_i}{\alpha_i} \right)^2 \right]^{\beta_i}} \quad i = 3, 4; \quad (2.12)$$

Couche multiplication : Multiplication des valeurs floues pour avoir les poids de chaque règle.

$$O_k^2 = w_k = \mu_{A_i}(x) \cdot \mu_{B_j}(y) \quad k = 1, 2, 3, 4; i = 1, 2; j = 1, 2; \quad (2.13)$$

Couche normalisation : Normalisation des poids.

$$O_i^3 = \bar{w}_i = \frac{w_i}{\sum_{k=1}^4 w_k} \quad i = 1, \dots, 4; \quad (2.14)$$

Couche défuzzification : Calcul des valeurs de défuzzification par les paramètres conséquents $\{p_i, q_i, r_i\}$ et les poids \bar{w}_i .

$$O_i^4 = \bar{w}_i f_i = \bar{w}_i (p_i x + q_i y + r_i) \quad i = 1, \dots, 4; \quad (2.15)$$

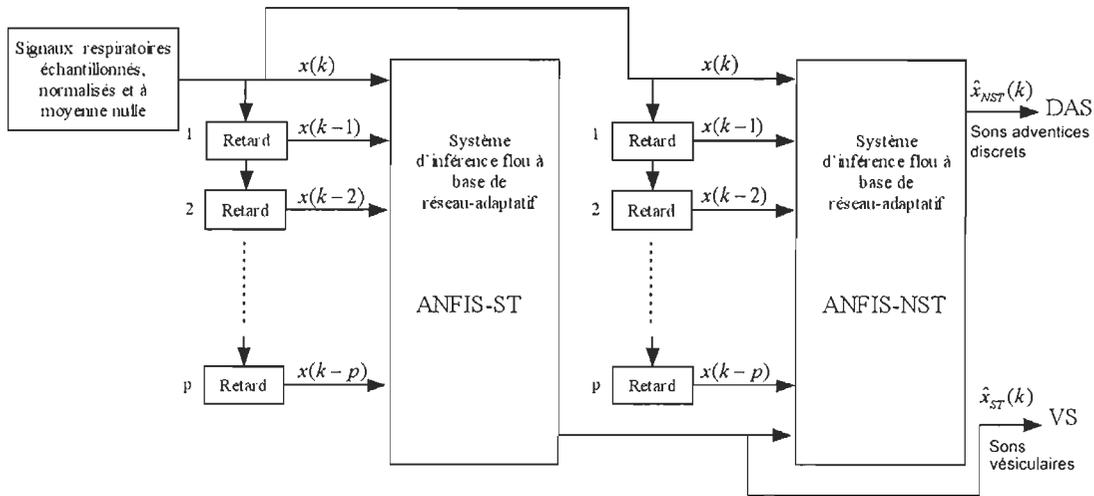


Figure 2.12: Schéma du filtre de séparation basé sur la logique floue GFST-NST [39].

Couche Agrégation : Calcul de la sortie finale par l'agrégation.

$$O^5 = \sum_{i=1}^4 \bar{w}_i f_i = \frac{\sum_i^4 w_i f_i}{\sum_i^4 w_i} \quad (2.16)$$

2.3.4 Filtre de séparation basé sur la logique floue GFST-NST

Le filtre GFST-NST est composé de deux systèmes ANFIS (voir la figure 2.12). Ces deux systèmes sont entraînés grâce à une base de données qui comporte les résultats de séparation obtenus par d'autres méthodes. On pourrait obtenir sensiblement les mêmes résultats qu'avec le filtre WTST-NST mais avec beaucoup moins de temps. Cependant, le filtre GFST-NST nécessite un entraînement séparé pour chaque type de son (crépitant fin et gros crépitant). Dans le travail de Tolia et al. [39], la base de données a été construite en créant un fichier audio par type de crépitant. En d'autres termes, nous avons besoin d'informations a priori sur le signal avant la séparation, afin de choisir le filtre approprié.

2.4 Extraction des crépitants par la transformée en ondelettes

La transformée en ondelettes a suscité, ces dernières années, un intérêt considérable et trouve des applications dans de nombreux domaines. Dans le cas des sons respiratoires, un filtre de séparation des crépitants de sons vésiculaires, basé sur la transformée en ondelettes a été proposé par Hadjileontiadis et al. [38]. Ce filtre de séparation présente d'excellentes performances. Avant d'aborder le principe de fonctionnement de ce filtre, nous ferons un petit rappel de la transformée en ondelettes et son implantation discrète.

2.4.1 Transformée en ondelettes

La transformée en ondelettes continue d'un signal $x(t)$ est définie par :

$$w(s, \tau) = \langle x(t), \psi_{s,\tau} \rangle = \frac{1}{\sqrt{s}} \int_{-\infty}^{+\infty} x(t) \psi_{s,\tau}^*(t) dt \quad (2.17)$$

où $(*)$ désigne l'opérateur conjugué, et $\psi_{s,\tau}(t)$ est une base orthogonale d'ondelettes, générée à partir d'une seule ondelette $\psi(t)$ qui s'appelle ondelette mère ou ondelette analysante :

$$\psi_{s,\tau}(t) = \frac{1}{\sqrt{s}} \psi \left(\frac{t - \tau}{s} \right) \quad (2.18)$$

où $s \in \mathbb{R}^+$ est le paramètre de dilatation caractérisant l'échelle du temps et $\tau \in \mathbb{R}$ est le paramètre de translation caractérisant l'emplacement sur l'axe du temps. L'ondelette mère doit être de carré intégrable et vérifier la condition d'admissibilité. Plusieurs types d'ondelettes sont proposés dans la littérature, par exemple l'ondelette de *Haar*, l'ondelette de *Morlet*, et l'ondelette de *Daubechies* que nous utiliserons dans ce projet.

Pour une implantation numérique, il est nécessaire de représenter de façon discrète les paramètres s et τ [50]. Ils sont, le plus souvent, définis par un format dyadique : $s = 2^j$ et $\tau = 2^j k$, $k \in \mathbb{Z}$. L'équation 2.17 devient :

$$w_j[k] = 2^{-\frac{j}{2}} \int_{-\infty}^{+\infty} x(t) \psi^*(2^{-j}t - k) dt \quad (2.19)$$

L'analyse et la synthèse par transformée en ondelettes consistent à utiliser deux ondelettes $\psi(t)$ et $\phi(t)$ de façon que l'on puisse écrire ^[51] :

$$x(t) = \sum_{j=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} w_j[k] \phi(2^{-j}t - k) \quad (2.20)$$

2.4.2 Analyse multirésolution

L'analyse multirésolution consiste à projeter un signal, $x(t) \in L^2(\mathbb{R})$, appartenant à un espace V_j sur un sous-espace V_{j+1} et un sous-espace W_{j+1} . V_{j+1} est sous-espace d'approximation et W_{j+1} est le sous-espace détails ^[52]. Ces sous-espaces ont un certain nombre de propriétés : ils forment une suite emboîtée, leur intersection est réduite à $\{\emptyset\}$ et leur réunion est dense dans $L^2(\mathbb{R})$. Il est démontré qu'il existe une fonction d'échelle $\varphi(t) \in L^2(\mathbb{R})$ qui engendre une base orthonormée de V_{j+1} et une ondelette $\psi(t) \in L^2(\mathbb{R})$ qui engendre une base orthonormée de W_{j+1} .

$$a_j(k) = \langle x(t), \varphi_{j,k} \rangle = \int_{-\infty}^{+\infty} x(t) \varphi_{j,k}^*(t) dt \quad a_j(k) \in V_j \quad (2.21)$$

$$d_j(k) = \langle x(t), \psi_{j,k} \rangle = \int_{-\infty}^{+\infty} x(t) \psi_{j,k}^*(t) dt \quad d_j(k) \in W_j \quad (2.22)$$

où les bases orthonormées sont définies par :

$$\varphi_{j,k}(t) = 2^{-\frac{j}{2}} \varphi(2^{-j}t - k) \quad j, k \in \mathbb{Z} \quad (2.23)$$

$$\psi_{j,k}(t) = 2^{-\frac{j}{2}} \psi(2^{-j}t - k) \quad j, k \in \mathbb{Z} \quad (2.24)$$

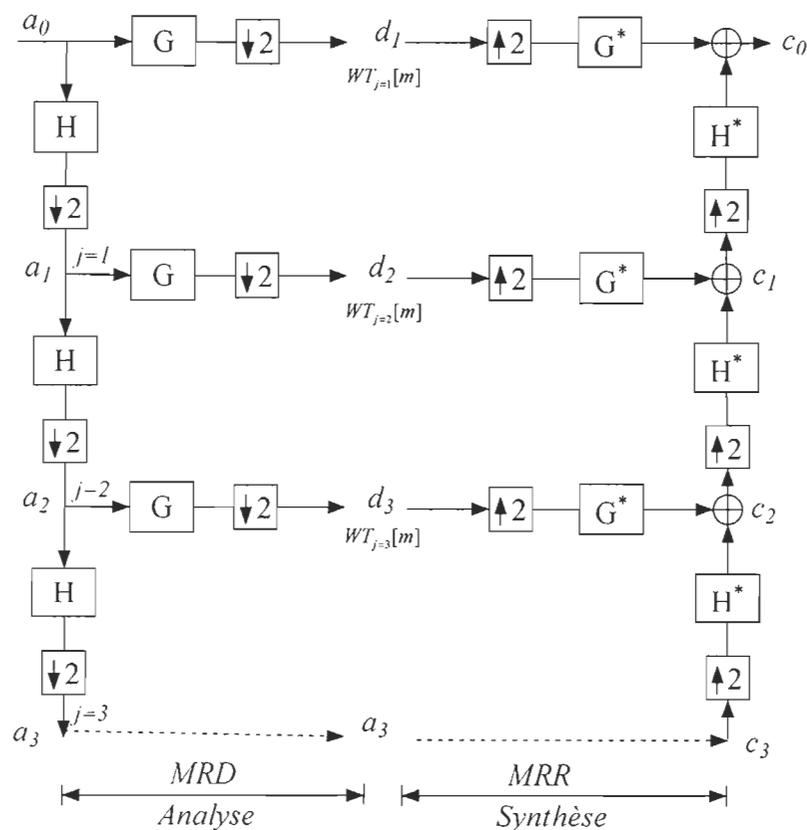


Figure 2.13: Schéma d'analyse MRD (*Multi-Resolution Decomposition*) et de reconstitution MRR (*Multi-Resolution Reconstruction*) de l'algorithme de Mallat [53].

Les sous-espaces V_j et W_j se caractérisent par :

$$V_{j-1} = V_j \oplus W_j \quad (2.25)$$

$$V_0 = W_1 \oplus \dots \oplus W_J \oplus V_J \quad (2.26)$$

L'analyse multirésolution utilisée dans ce document est basée sur l'algorithme de Mallat [53]. Cet algorithme est basé sur la définition de deux filtres H et G , dont les réponses impulsionnelles sont reliées aux fonctions $\varphi(t)$ et $\psi(t)$. La figure 2.13 présente le schéma de décomposition MRD (*Multiresolution Decomposition*) et de reconstitution MRR (*Multiresolution Reconstruction*) de cet algorithme.

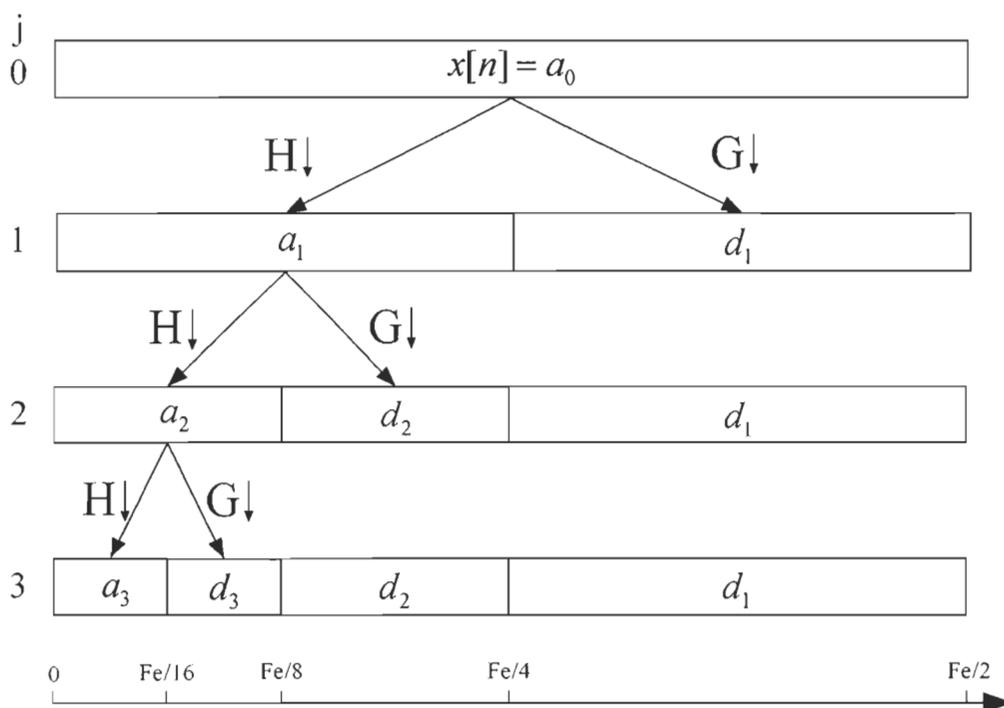


Figure 2.14: Répartition fréquentielle par transformée en ondelettes.

Les filtres H et G constituent respectivement, un filtre passe-bas et un filtre passe-haut ayant la même fréquence de coupure (leurs bandes de fréquence sont complémentaires), appelés aussi filtres à miroirs quadratiques (QMF) ^[51]. Durant l'analyse (décomposition) du signal, les sorties des filtres H et G sont sous-échantillonnées par un facteur 2 à chaque échelle j . Chaque approximation a_j est décomposée par ces deux filtres pour donner une nouvelle approximation a_{j+1} et un nouveau détail d_{j+1} . La synthèse (reconstitution) du signal est un procédé similaire à l'analyse du signal mais opérant en sens inverse.

Dans le domaine des fréquences, le signal est décomposé en plusieurs sous-bandes de fréquence (voir la figure 2.14). C'est évident que la meilleure résolution fréquentielle est obtenue en basses fréquences, et inversement, la meilleure résolution temporelle est obtenue en hautes fréquences.

2.4.3 Filtre de séparation basé sur la transformée en ondelettes WTST-NST

Le filtre WTST-NST est une approche proposée par Hadjileontiadis et al. [38] pour séparer les sons adventices discontinus des sons vésiculaires. Cette méthode de séparation est basée sur l'analyse multirésolution par ondelettes (figure 2.13). Elle est décrite par l'algorithme suivant :

1. *Initialisation :*

$x[n]$: un segment de N échantillons ($N = 2^M$) d'un signal respiratoire,

F_{adj} : un facteur multiplicatif d'ajustement,

ε : une constante positive de très faible valeur,

j : l'indice de l'échelle ($j = 1, \dots, M$) avec $M = \log_2(N)$.

2. *Attribuer :* $f[n] = x[n]$

3. *Boucle pour chaque itération :* $i = 1, \dots, L$

{

Calcul des coefficients $WT^i = MRD(f[n])_j$ *échelles*

Sauvegarder deux copies de WT^i *en* WT^{iC} *et* WT^{iR} . *Autrement dit,* $WT^{iC} =$

$WT^{iR} = WT^i$

Boucle pour chaque échelle d'ondelettes : $j=1, \dots, M$.

{

Calculer les écarts-types σ_j^i

Calculer les seuils $THR_j^i = \sigma_j^i \cdot F_{adj}$

Boucle des coefficients d'ondelettes à chaque échelle : $m = 1, \dots, \frac{N}{2^j}$.

{

Comparer WT_j^i *et* THR_j^i

Si $|WT_j^i| < THR_j^i$ *alors* $WT_j^{iC}[m] = 0$ *sinon* $WT_j^{iR}[m] = 0$

} fin boucle m

} fin boucle j

Calculer $C^i[n] = MRR(WT^{iC})_j$ *échelles*

Calculer $R^i[n] = MRR(WT^{iR})_j$ échelles

Calculer le critère $STC = |E\{R_{i-1}^2[n]\} - E\{R_i^2[n]\}|$

Compare STC à ε :

Si $STC \geq \varepsilon$ alors $f(n) = R^i[n]$ sinon $i = L$ et quitter la boucle

} fin boucle i

4. Calculer les sons adventices discontinus DSA ou $x_d[n]$:

$$DAS[n] = x_d[n] = \sum_{i=1}^L C_i[n]$$

5. Calculer les sons vésiculaires purs PVS ou $x_c[n]$:

$$PVS[n] = x_c[n] = R_L[n]$$

L'organigramme correspondant est présenté à la figure 2.15 de façon très explicite. Les signaux $x_d[x]$ et $x_c[x]$ représentent respectivement les sons adventices discontinus et les sons vésiculaires.

Malgré qu'elle présente d'excellentes performances de séparation, cette technique est très coûteuse en temps de calcul, ce qui rend très difficile son implantation en temps réel. En effet, ce handicap est causé par deux facteurs :

1. L'analyse par ondelettes se fait sur toutes les échelles possibles. Par exemple, si la fenêtre du signal analysé contient $N=1024$ échantillons, alors le nombre d'échelles $M = \log_2(N) = 10$. Autrement dit, $j = 1, \dots, 10$.
2. L'algorithme est itératif. Il est donc répété L fois.

2.5 Extraction des crépitants par les paquets d'ondelettes

Dans cette section, nous présentons une nouvelle technique d'extraction des crépitants des sons respiratoires, que nous avons proposée récemment ^[40]. Cette méthode, basée sur les paquets d'ondelettes, exploite les caractéristiques temps-fréquence du signal. De plus, la transformée par paquets d'ondelettes nous offre la possibilité de profiter du choix immense de toutes les bases possibles pour trouver une segmentation optimale, qui soit la mieux adaptée au problème de détection des transitoires. Ainsi, la transformée en paquets d'ondelettes fournit une segmentation dynamique selon les caractéristiques du signal.

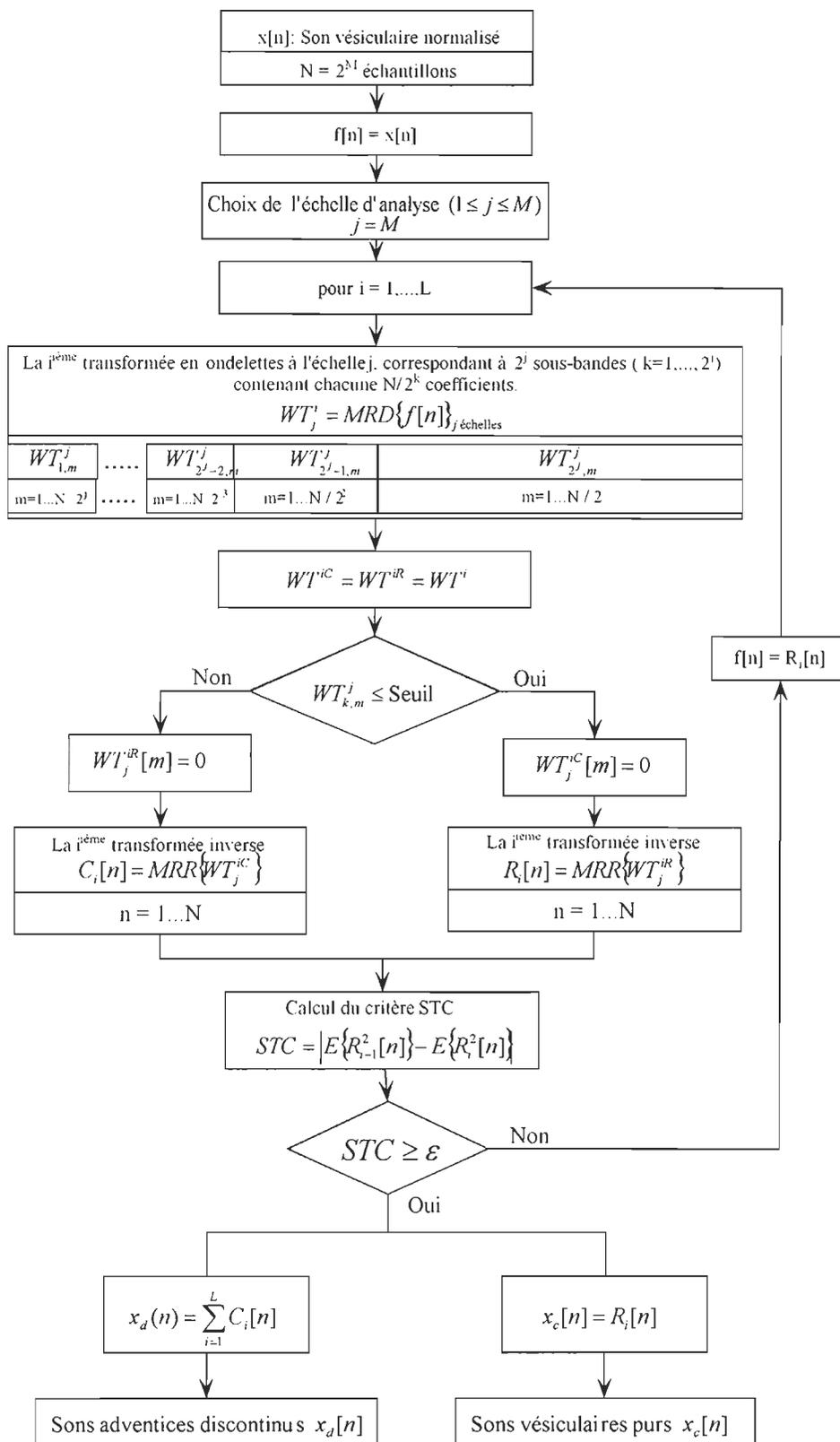


Figure 2.15: Schéma du filtre de séparation basé sur la transformée en ondelettes WTST-NST [38].

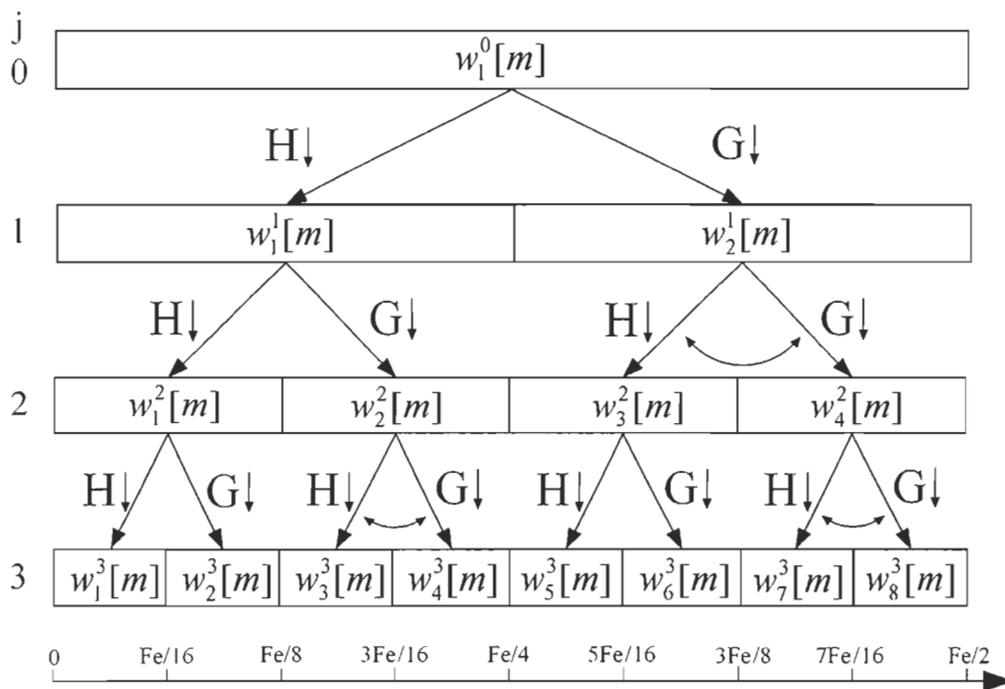


Figure 2.16: Répartition fréquentielle par paquets d'ondelettes.

2.5.1 Transformée par paquets d'ondelettes

La transformée par paquets d'ondelettes, proposée par Coifman et al ^[54], se présente comme une généralisation de la transformée en ondelettes. Pour une résolution donnée, les filtres H et G sont appliqués simultanément à l'approximation (a_j) et au détail (d_j). La figure 2.16 présente un exemple d'analyse par paquets d'ondelettes, où l'arbre de décomposition dyadique est complet.

D'après les travaux de Wickerhauser ^[55], les sous-bandes de l'arbre des paquets d'ondelettes ne sont pas organisées dans l'ordre naturel croissant des fréquences. Pour avoir une distribution régulière sur l'axe fréquentiel, il faut réarranger certains filtres H et G lors de la décomposition du signal (voir la figure 2.16).

Pour simplifier, nous utiliserons la même notation pour tous les coefficients. En effet, $w_k^j[m]$ représente le $m^{\text{ième}}$ coefficient de la $k^{\text{ième}}$ sous-bande de l'échelle j .

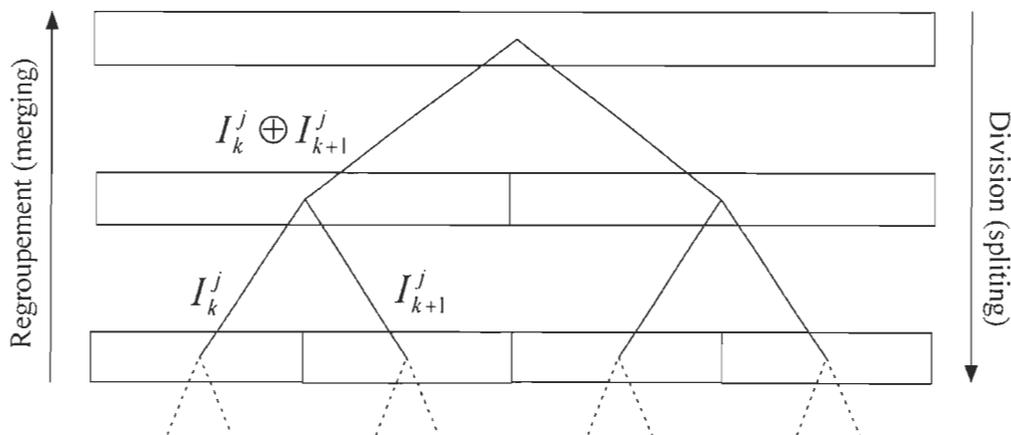


Figure 2.17: Deux techniques de recherche de la meilleure base.

2.5.2 Recherche de la meilleure base

La problématique de la meilleure base vise à chercher les composantes les plus significatives du signal. Elle est utilisée, le plus souvent, pour l'extraction des caractéristiques et la compression des données.

La transformée par parquets d'ondelettes décompose le signal d'entrée en plusieurs sous-bandes dont la largeur dépend de la profondeur j de l'arbre de décomposition. La base optimale "la meilleure" peut être interprétée comme la meilleure segmentation de l'axe des fréquences ^[51]. Deux techniques sont utilisées pour trouver la meilleure base : la méthode de division (*splitting*) et la méthode de regroupement (*merging*). Par exemple, la méthode de regroupement cherchera la meilleure base en partant des feuilles en se dirigeant vers la racine de l'arbre, ou vice versa pour la méthode de division (voir la figure 2.17).

La meilleure base est déterminée en minimisant une fonction de coût (critère d'optimisation). Le choix de la fonction de coût dépend de l'application en question. Le critère le plus utilisé dans la littérature est *l'entropie* du spectre ^[56] :

$$H(I_k^j) = - \sum_m |w_k^j[m]|^2 \log |w_k^j[m]|^2 \quad (2.27)$$

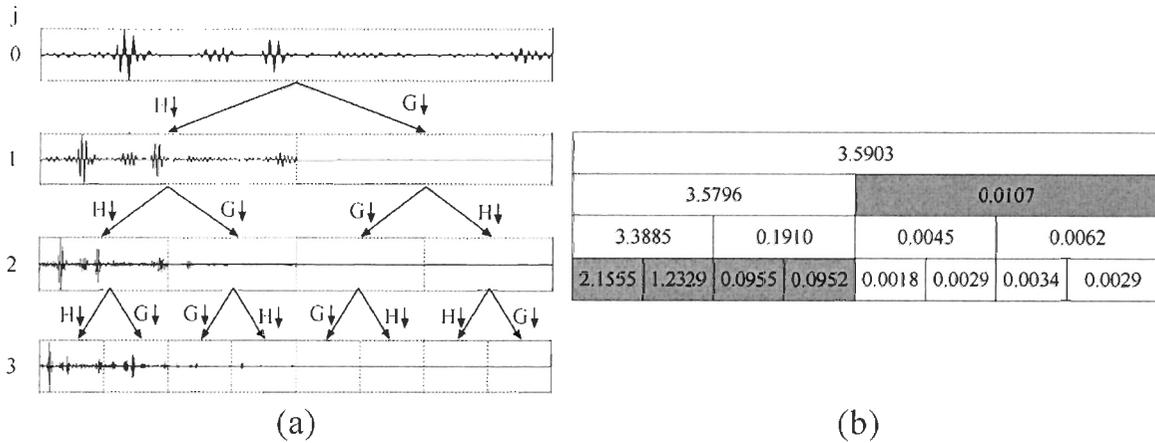


Figure 2.18: Exemple d'analyse de crépitants par paquets d'ondelettes : (a) les coefficients de chaque sous-bande; (b) le coût "entropie" de chaque sous-bande.

Dans la technique ascendante, on commence par une segmentation arbitraire de la bande de fréquence en intervalles dyadiques de taille réduite, pour que l'on puisse prendre en compte les détails les plus fins apparaissant dans le signal. L'algorithme consiste, alors, à remplacer les intervalles contigus \mathbb{I}_k^j et \mathbb{I}_{k+1}^j (intervalles fils) par l'union des deux $\mathbb{I}_k^j \oplus \mathbb{I}_{k+1}^j$ (intervalle père), selon un critère d'optimisation ^[51] :

$$H(\mathbb{I}_k^j \oplus \mathbb{I}_{k+1}^j) \lesssim H(\mathbb{I}_k^j) + H(\mathbb{I}_{k+1}^j) \quad (2.28)$$

La figure 2.18 représente un exemple de recherche de la meilleure base. En effet, la figure 2.18-a représente la forme du signal (les coefficients $w_k^j[m]$) dans chaque sous-bande. En revanche, la figure 2.18-b représente le coût "entropie" de chaque sous-bande par un nombre. La meilleure base, dans cet exemple, correspond aux sous-bandes $[w_1^3, w_2^3, w_3^3, w_4^3, w_2^1]$ et est représentée en gris dans la figure 2.18-b.

2.5.3 Filtre de séparation basé sur les paquets en ondelettes WPST–NST

En analysant les caractéristiques temporelles et fréquentielles des sons adventices discontinus (figure 1.4), on constate que le crépitant se caractérise par une amplitude qui augmente remarquablement durant une période courte et une énergie qui s'étend jusqu'à

1500 Hz environ. Les deux caractéristiques démontrent que le crépitant est un signal transitoire et explosif.

Dans cette section, nous présentons en détails l'algorithme à base de paquets d'ondelettes (WPST-NST) que nous avons proposé pour séparer les crépitants des sons adventices [40]. Le principe du filtre WPST-NST repose sur l'idée d'utiliser l'amplitude et la distribution d'énergie du signal comme indicateur de présence d'un crépitant.

L'algorithme du filtre WPST-NST est décrit comme suit :

1. Diviser le signal en segments de taille L et entrelacés de 75%.
2. Décomposer chacun des segments par les parquets d'ondelettes (WPT).
3. Marquer les coefficients $w_k^j[m]$ dont l'amplitude dépasse le seuil ST_k^j défini par :

$$ST_k^j = P_1 * \sigma_k^j \quad (2.29)$$

où σ_k^j est l'écart-type des coefficients relatifs à la $k^{ième}$ sous-bande du niveau j . La constante P_1 est fixée expérimentalement à 0.75.

Ainsi, on définit le marquage $M_k^j[m]$ d'un coefficient $w_k^j[m]$ comme suit :

$$M_k^j[m] = \begin{cases} 1 & \text{si } |w_k^j[m]| \geq ST_k^j \\ 0 & \text{sinon} \end{cases} \quad (2.30)$$

4. Quantifier le nombre de marquage d'un coefficient $w_k^j[m]$ pour l'ensemble des sous-bandes k d'un même niveau j :

$$N^j[m] = \sum_{k=1}^{2^j} M_k^j[m] \quad (2.31)$$

On définit un deuxième seuil $S2_k^j$ relié à la moyenne des nombres de marquage par sous-bande :

$$S2_k^j = P_2 * \frac{1}{L/2^j} \sum_{m=1}^{L/2^j} N^j[m] \quad (2.32)$$

La constante P_2 est fixée expérimentalement à 2.

5. Séparer les coefficients $w_k^j[m]$ en deux composantes : stationnaires $S_ST_k^j[m]$ et non-stationnaires $S_NST_k^j[m]$ selon les valeurs des seuils ST_k^j et $S2_k^j$.

Initialisation $S_NST_k^j[m] := 0; S_NST_k^j[m] := 0;$

si $|w_k^j[m]| \geq ST_k^j$ et $N^j[m] \geq S2^j$, alors $S_NST_k^j[m] = w_k^j[m];$

sinon $S_ST_k^j[m] = w_k^j[m].$

6. Chercher la meilleure base de décomposition pour les coefficients $w_k^j[m]$.
7. Chaque segment est divisé par le nombre de chevauchement. Reconstruire les signaux stationnaire et non-stationnaire par transformation inverse (IWPT) des coefficients correspondants $S_ST_k^j[m]$ et $S_NST_k^j[m]$.

L'organigramme correspondant au filtre WPST-NST est présenté à la figure 2.19 de façon très explicite. Les signaux $x_d[x]$ et $x_c[x]$ représentent respectivement les sons adventices discontinus et les sons vésiculaires.

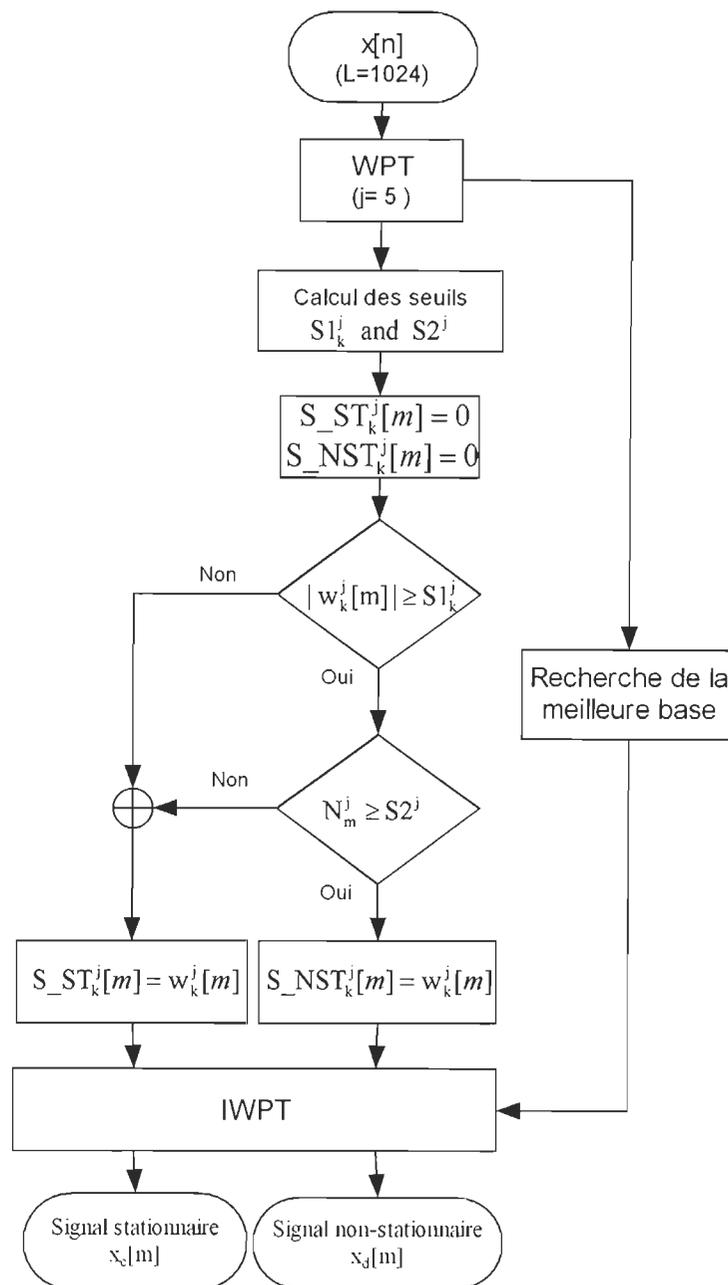


Figure 2.19: Schéma du filtre de séparation basé sur les paquets en ondelettes WPST-NST [40].

CHAPITRE III

IDENTIFICATION DES CRÉPITANTS

Dans ce chapitre, nous allons présenter en détails le module d'identification qui suit le filtre de séparation de la figure 1.6. Ce module opère en trois étapes principales : 1) un filtre de débruitage est appliqué pour supprimer le bruit résiduel dans le signal non-stationnaire, 2) un détecteur basé sur la dimension fractale localise la position du crépitant, et 3) les crépitants sont classifiés en gros crépitants et en crépitants fins selon leurs caractéristiques.

3.1 Principe

L'analyse des crépitants fournit des informations très utiles lors du diagnostic des maladies pulmonaires. Le nombre des crépitants dans un cycle respiratoire reflète la sévérité de la maladie, la durée et l'instant d'apparition des crépitants peuvent varier selon le type de maladies ^[28].

Après le filtre de séparation, le signal respiratoire est divisé en deux signaux : un signal stationnaire (ST-Signal) et un signal non-stationnaire (NST-Signal). Le ST-Signal contient les sons vésiculaires et les sons adventices continus; alors que le NST-Signal contient les crépitants auxquels nous nous intéressons dans ce chapitre. Par le procédé de séparation, nous minimiserons l'influence du signal stationnaire (ST-Signal) durant l'analyse du signal non-stationnaire (NST-Signal).

L'identification des crépitants consiste à localiser les crépitants individuellement puis les classifier en crépitants fins ou en gros crépitants. Les signaux analysés dans ce projet

proviennent de la sortie de notre filtre de séparation WPST-NST. Comme pré-traitement, un filtre de débruitage a été proposé pour mettre les échantillons non-crépitanants à zéro dans le signal non-stationnaire NST-Signal. L'étape de détection cherche l'instant d'apparition de chacun des crépitants puis les crépitants sont extraits individuellement. Enfin, nous allons définir les paramètres caractéristiques des crépitants en vue de les classifier.

3.2 Filtre de débruitage

L'amplitude des crépitants n'est pas constante le long d'un signal respiratoire et peut varier dans une plage relativement étendue. Après l'étape de séparation, il y a encore du bruit résiduel dans le signal non-stationnaire NST-Signal. Parfois le crépitant est trop faible pour se distinguer du bruit uniquement par son énergie ou par son amplitude. Dans ce projet, nous proposons un filtre de débruitage très efficace pour réduire les bruits résiduels dans le NST-Signal. Le filtre proposé est caractérisé par un seuil adaptatif dépendant du signal stationnaire ST-Signal. Il est décrit par l'algorithme suivant :

1. Générer un ensemble d'index vide $A = \emptyset$ et un ensemble d'index plein $\Omega = \{1, 2, \dots, N\}$, où N est égale au nombre d'échantillons du signal. Initialiser $S_NST_f[n] = S_NST[n]$.
2. Appliquer un fenêtrage entrelacé ($\alpha = 75\%$) simultanément aux signaux $S_ST[n]$ et $S_NST[n]$. La taille de la fenêtre L est approximativement égale à la moitié de la longueur d'un crépitant. À la fréquence d'échantillonnage de 5 kHz, nous définissons $L=32$.
3. Boucle pour chaque fenêtre du signal $k = 1, 2, \dots, K$:
 - {
 - Calculer la somme des amplitudes absolues du signal dans chaque fenêtre :

$$A_{ST}[k] = \sum_{n=(k-1)(1-\alpha)L+1}^{(k-1)(1-\alpha)L+L} |S_ST[n]| \quad (3.1)$$

$$A_{NST}[k] = \sum_{n=(k-1)(1-\alpha)L+1}^{(k-1)(1-\alpha)L+L} |S_NST[n]| \quad (3.2)$$

où n est l'index d'échantillonnage.

- Si $A_{NST}[k] > P * A_{ST}[k]$, alors réunir les indexes d'échantillonnage du segment k avec l'ensemble A , où P est une constante fixée à 0.8 expérimentalement :

$$A = A \cup \{(k-1)(1-\alpha)L + 1, (k-1)(1-\alpha)L + 2, \dots, (k-1)(1-\alpha)L + L\} \quad (3.3)$$

} fin de boucle k

4. Les échantillons qui n'appartiennent pas à l'ensemble A sont mis à zéro.

$$S_{NST_f}(\overline{\Omega \cap A}) = 0 \quad (3.4)$$

Ce filtre préserve tous les crépitants qu'ils soient faibles ou forts, tout en supprimant efficacement le bruit.

3.3 Détection des crépitants

La détection des crépitants sert à les localiser individuellement pour les classifier ensuite. En général, la détection est réalisée à l'aide d'un seuil qui permet de distinguer les crépitants du son stationnaire. Plusieurs solutions, traitant cette problématique, ont déjà été proposées dans la littérature : la détection selon la première dérivée du signal ^[57], la détection selon la "stationnarité" de la matrice sonagramme ^[58] et la détection selon le paramètre de la dimension fractale (FD) ^[41]. Dans notre cas, la problématique de détection des crépitants est rendue plus facile par le filtre de séparation. Nous considérons que le NST-Signal, après débruitage, contient seulement des crépitants.

Les crépitants peuvent être faibles ou forts en amplitude, parfois certains d'entre eux apparaissent consécutivement. Le problème est que, si le seuil de détection est surestimé, nous pourrions perdre ces crépitants faibles ; par contre, si le seuil est sous-estimé, ces crépitants consécutifs pourraient être mal-détectés et pris pour un seul crépitant. Alors, il est difficile de choisir un seuil universel qui soit approprié dans toutes les situations. Au lieu de prendre un seuil fixe, nous pourrions utiliser un seuil adaptatif relié à l'amplitude

du crépitant.

Normalement, un crépitant comprend 4 à 10 déflexions positives ou négatives, l'amplitude de chaque déflexion augmente graduellement puis diminue après avoir atteint son sommet. Si on divise un crépitant en deux parties par rapport à son amplitude absolue maximale, l'enveloppe du crépitant serait donc monotone. La méthode proposée ^[34] pour la détection de crépitants est basée sur cette propriété.

3.3.1 Choix du paramètre de la détection

Le paramètre utilisé pour la détection doit refléter le changement de l'amplitude du crépitant tout en ayant une forme d'onde lisse pour éviter les erreurs de jugement. Trois paramètres sont comparés dans cette section : la dimension fractale, l'enveloppe de l'amplitude absolue et l'enveloppe de l'énergie. Pour comparer la performance de ces paramètres, nous prenons l'exemple illustré à la figure 3.1-a. Cet exemple se compose de 7 crépitants, 3 d'entre eux sont entrelacés (4^{ième}, 5^{ième} et 6^{ième}).

L'enveloppe de l'amplitude est obtenue par un filtre passe-bas (ayant une fréquence coupure de 20 Hz dans cet exemple). Ce paramètre est capable de discriminer les trois crépitants entrelacés, mais en même temps plusieurs sous-pics sont présents au milieu de plusieurs crépitants, pouvant causer des fausses détections (voir la figure 3.1-c).

L'enveloppe de l'énergie est calculée par la transformée de Fourier à court terme (STFT) ^[58]. Le signal filtré $S_{NST_f}[n]$ est divisé en segments entrelacés avant de lui appliquer la transformée de Fourier. La longueur de la fenêtre est fixée à 12.8 ms avec 99% de chevauchement. Dans chaque segment, l'énergie du signal est affectée au point de la mi-fenêtre (voir la figure 3.1-d). Pour ce qui est des trois crépitants entrelacés, l'enveloppe de l'énergie du signal sépare les deux premiers mais le troisième qui est plus faible est couvert par le deuxième crépitant, donc, ce paramètre n'est pas assez précis pour détecter les crépitants faibles.

La détection des crépitants par dimension fractale a été initialement proposée par

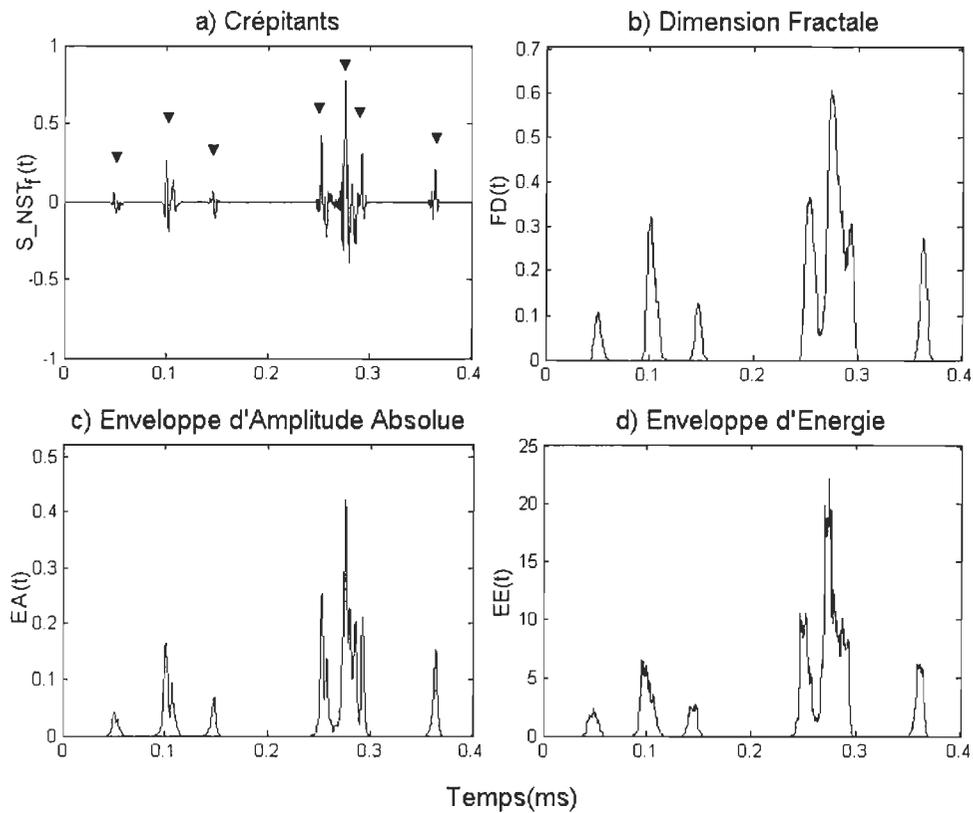


Figure 3.1: Comparaison de trois paramètres utilisés dans la détection des crépitants. a) Sortie filtré $S_NST_f[n]$ contenant (07) crépitants identifiés par les symboles "▼", b) Dimension Fractale (FD), c) Enveloppe de l'amplitude absolue (EA), d) Enveloppe de l'énergie (EE).

Hadjileontiadis et Rekanos ^[41]. Le paramètre FD présente la meilleure performance. L'ensemble des 7 crépitants est clairement séparé (voir la figure 3.1-b). Nous allons introduire la notion de la dimension fractale dans la section suivante avant de présenter l'algorithme de la détection des crépitants.

3.3.2 Notion de la dimension fractale

Dans le sens commun, la notion de "dimension" est une mesure qui reflète la taille d'un objet, par exemple la longueur, la largeur, la profondeur, etc. En mathématiques, les dimensions sont les paramètres qui décrivent la position et l'état d'un objet dans un

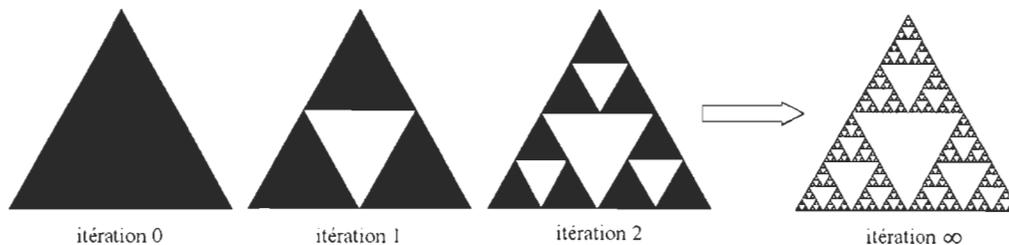


Figure 3.2: Construction du triangle de *Sierpinski*.

espace conceptuel. La dimension d'un espace est égale au nombre total de paramètres qui servent à définir tous les objets possibles dans cet espace ^[59].

Selon les différents types d'espaces, les définitions de dimension sont nombreuses. L'espace euclidien est une conception plus élémentaire, il tire son nom du mathématicien grec *Euclide*. Dans l'espace euclidien, le point est de dimension 0, la courbe plane est de dimension 1, la surface est de dimension 2, le volume est de dimension 3. Cela s'appelle la dimension euclidienne (ou dimension topologique) et se caractérise par un nombre entier non-négatif.

Mais la définition de la dimension euclidienne est insuffisante pour décrire les objets irréguliers ou morcelés, comme les "fractals". Les objets fractals sont facilement observables dans la nature, par exemple les nuages, les montagnes, le brocoli, etc. Le plus souvent, ils possèdent des caractéristiques "auto-similaires" et pourraient être défini par une définition récursive ^[60]. Prenons l'exemple du triangle de *Sierpinski*, il se construit de manière itérative. Le principe est de partir, à l'itération $n = 0$, d'un triangle équilatéral plein, on efface le triangle du milieu dont les sommets sont les milieux des côtés du triangle initial. On obtient ainsi, à l'itération $n = 1$, trois petit triangles pleins ayant un côté deux fois plus petit. Cette opération est répétée une infinité de fois (voir la figure 3.2). Quand $n \rightarrow \infty$, sa surface est nulle et son périmètre est infini. Le triangle de *Sierpinski* est une ligne universelle ayant une dimension euclidienne de 1.

Comparée à la dimension euclidienne, la dimension fractale est un nombre réel non-négatif. Elle est calculée le plus souvent par la définition de la dimension *Hausdorff-Besicovitch*. Supposons qu'un objet est segmenté en N pièces identiques, le rapport d'ho-

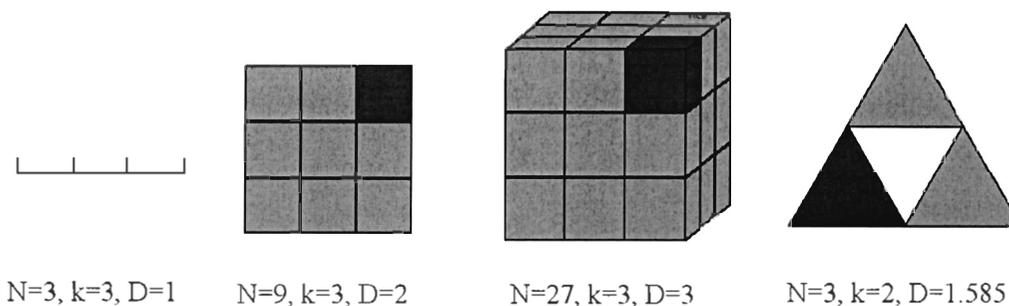


Figure 3.3: Dimension fractale calculée pour quatre objets. $D_h = \log(N)/\log(k)$, où N est le nombre de pièces identiques et k est le rapport d'homothétie (ou le taux de réduction)

mothétie k est égale au taux de réduction. Alors, la dimension *Hausdorff-Besicovitch* de cet objet est définie par :

$$D_h = \lim_{k \rightarrow \infty} \frac{\log(N)}{\log(k)} \quad (3.5)$$

Par exemple, la figure 3.3 donne un exemple de la dimension *Hausdorff-Besicovitch* de quatre objets : une ligne, un carré, un cube et un triangle *Sierpinski*. Parce que les quatre objets ont un caractère auto-similaire, leurs dimensions *Hausdorff-Besicovitch* sont constantes quelle que soit l'échelle que l'on choisit. Dans le cas du triangle de *Sierpinski*, sa dimension de *Hausdorff-Besicovitch* ($D_h = \log(3)/\log(2) = 1.585$) est supérieure à sa dimension euclidienne ($D_e = 1$). En effet, à chaque itération, un triangle donne trois nouveaux triangles $N = 3$, dont les côtés sont deux fois plus petits $k = 2$. La dimension fractale reflète la complexité d'un objet auto-similaire, ce qui permet au mathématicien français, *Benoît Mandelbrot* d'énoncer : Un objet est fractal si sa dimension de *Hausdorff-Besicovitch* est strictement supérieure à sa dimension topologique (sauf pour les courbes spatiales) [61].

On pourrait aussi utiliser la dimension fractale pour évaluer la complexité d'une courbe plane, dont l'abscise devrait augmenter monotonement. On considère une courbe $x[n]$ de N points, obtenus par échantillonnage uniforme de son abscise (le temps en général). Selon

l'algorithme de *Katz* ^[62], sa dimension fractale est définie par :

$$D_k = \frac{\log(N')}{\log(\frac{d}{L}) + \log(N')} \quad (3.6)$$

où $N' = N - 1$ est le nombre de segments de l'onde, d est le diamètre ou l'étendu de la courbe et L est la longueur de la courbe. Les paramètres d et L sont définis par :

$$d = \max_j \{dist(1, j)\} \quad (3.7)$$

$$L = \sum_{i=1}^{N'} dist(i, i+1) \quad (3.8)$$

où $dist(i, j)$ est la distance entre les points i et j de la courbe. Par cette définition, la courbe droite a une dimension fractale $D_k = 1$; la courbe aléatoire a une dimension $D_k = 1.15$ approximativement ; Pour la courbe la plus compliquée, sa dimension D_k s'approche de 1.5 ^[62].

Pour notre projet, nous utilisons un autre algorithme relativement simple proposé par *Sevcik* ^[63]. Il consiste à projeter l'abscisse et l'ordonnée dans un carré unitaire par une transformation linéaire. Le résultat est plus précis, surtout lorsque le nombre d'échantillons N est très grand.

$$D_s = 1 + \frac{\ln(L)}{\ln(2 * N')} \quad (3.9)$$

3.3.3 Détection par le paramètre "Dimension Fractale"

Après débruitage, le signal $S_NST_f[n]$ est segmenté en fenêtres entrelacées, la longueur de la fenêtre (N) est approximativement égale à la moitié de la longueur d'un crépitan ($N/Fs = 32/5000 \text{ Hz} = 6.4 \text{ ms}$). Le taux d'entrelacement est de 99%. La dimension fractale est calculée selon l'algorithme de *Sevcik* puis affectée à l'échantillon de la mi-fenêtre.

$$D_s[n] = 1 + \frac{\ln(L)}{\ln(2 * N')} \quad (3.10)$$

Dans le diagramme de la dimension fractale, chaque crépissant est représenté par un pic distinct (figure 3.1-b). Pour éviter l'influence des petites variations de cette courbe, une méthode statistique est développée afin de prédire le changement (augmentation ou diminution) de la courbe FD. Illustré à la figure 3.4, cet algorithme est décrit par les étapes suivantes :

1. *Enlever les petites variations de la courbe de la dimension fractale par :*

$$D'_s[n] = \begin{cases} D_s[n] - 1 - \varepsilon & \text{si } D_s[n] > 1 + \varepsilon \\ 0 & \text{sinon} \end{cases} \quad (3.11)$$

où ε est une constante de faible valeur fixée à 0.01. La valeur minimale de $D'_s[n]$ est 0. Les échantillons non-nuls sont considérés comme crépitants.

2. *Définir trois états de la courbe FD' : -1="décroissante", 0="constante" et 1="croissante".*
3. *Définir une fonction $X[n]$ égale à la dérivée de la courbe de la dimension fractale pour la quantifier ensuite à -1, 0 et 1 afin d'obtenir $U[n]$. Les trois valeurs de $U[n]$ correspondent à trois états de la courbe. On dit que $U[n]$ représente l'état de la courbe dans une période courte (de l'instant $n-1$ à n).*

$$X[n] = D'_s[n] - D'_s[n-1] \quad (3.12)$$

$$U[n] = \begin{cases} 1 & X[n] > 0 \\ 0 & X[n] = 0 \\ -1 & X[n] < 0 \end{cases} \quad (3.13)$$

4. *Définir une variable $E[n] \in [-1, 0, 1]$ représentant l'état instantané de la courbe et l'initialiser à zéro ($E[0] = 0$).*
5. *Définir une variable $T[n] \in [-1, 0, 1]$ qui représente la tendance du changement de la*

courbe $D'_s[n]$ à l'instant n :

$$T[n] = \begin{cases} -1 & \text{if } \sum_{i=n+1}^{n+P} X[n] < 0 \\ 0 & \text{if } \sum_{i=n+1}^{n+P} X[n] = 0 \\ 1 & \text{if } \sum_{i=n+1}^{n+P} X[n] > 0 \end{cases} \quad (3.14)$$

où $P = \frac{N}{2} + 1$ et N est la longueur de la fenêtre.

6. Initialiser deux ensembles $DSB = DSE = \emptyset$ pour enregistrer le début et la fin de chaque crépitan.

La figure 3.4 représente l'algorithme automatique d'énumération des crépitants capable de trouver le début et la fin de chaque crépitan individuel, qui les enregistre dans deux ensembles "DSB" et "DSE". Cet algorithme présente l'avantage de chercher les frontières des crépitants qu'ils soient faibles ou entrelacés.

Cet algorithme de détection fournit une précision de presque 100%. À l'aide des deux ensembles "DSB" et "DSE", on peut facilement calculer le nombre et la durée des crépitants. La figure 3.5 donne un exemple de détection par la dimension fractale D_s , où le signal testé ne contient que des crépitants parce qu'il a été préalablement traité par les filtres de séparation et de débruitage. Selon le résultat obtenu dans cet exemple, le taux de détection est de 100 %.

3.4 Classification des crépitants

Nous avons vu à la section 1.1.3 que le type de crépitan nous informe sur la nature de la maladie. Par exemple, un gros crépitan est présent dans les cas de broncho-pneumonie et de bronchiectasies, un crépitan fin est présent dans les cas de fibrose interstitielle et de pneumonie [28]. Dans cette section, nous allons comparer trois types de paramètres caractéristiques en vue de classifier les crépitants : l'analyse temps-fréquence [57,64,65], le modèle de Prony [66] et la transformée en ondelettes continue [67].

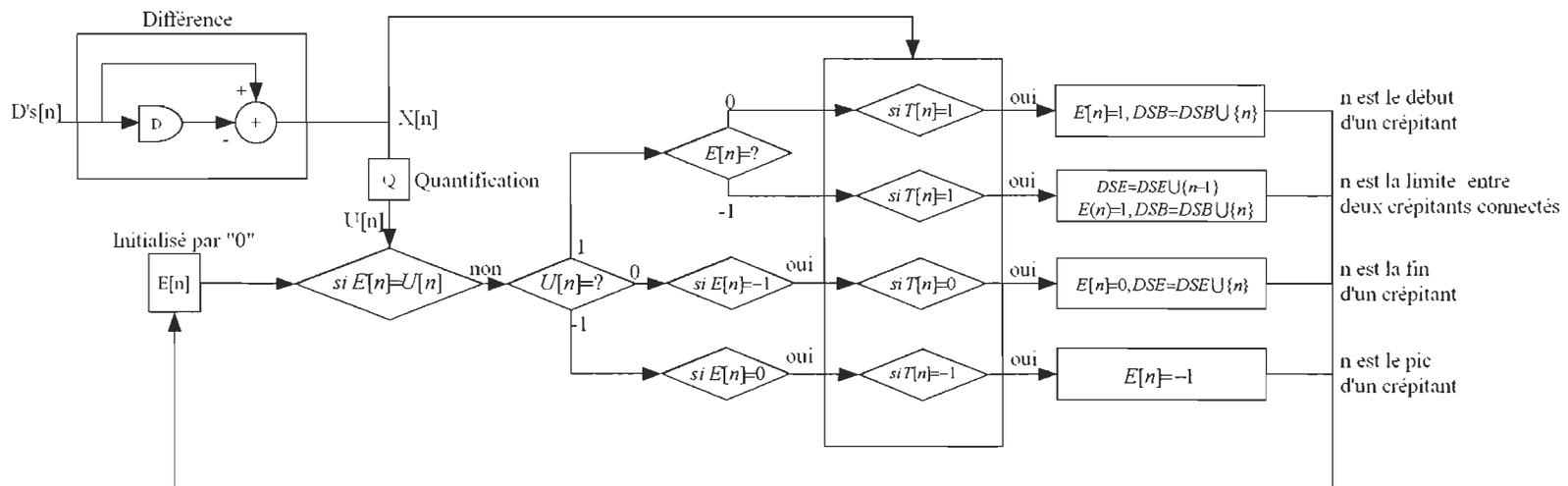


Figure 3.4: Algorithme automatique d'énumération des crépitations.

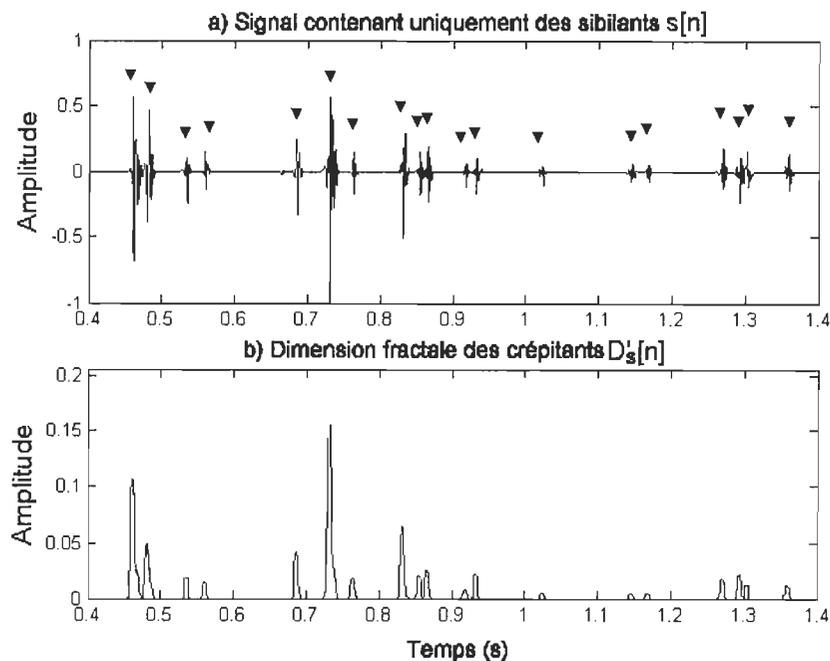


Figure 3.5: Exemple de détection des crépitations par dimension fractale. a) Le signal testé contient uniquement des crépitations parce qu'il a été préalablement traité par les filtres de séparation et de débruitage, où les crépitations sont identifiées par les marques (▼). b) La dimension fractale permet de détecter les crépitations faibles ou entrelacés.

3.4.1 Caractérisation temps-fréquence

La caractérisation temps-fréquence est une approche traditionnelle d'analyse de signaux respiratoires. Plusieurs mesures dans le domaine du temps ont été proposées pour différencier les deux types de crépitations : la durée de la déflexion initiale (IDW), la durée des deux premiers cycles (2CD), la durée de la déflexion maximale (MDW), la durée de la déflexion totale (TDW). Selon la proposition de l'ATS (*American Thoracic Society*), les durées moyennes de IDW et 2CD du crépitation fin sont respectivement de 0.7 ms et 5 ms, et celles des gros crépitations sont respectivement de 1.5 ms et 10 ms.

Les crépitations sont extraits individuellement dans l'étape de détection, mais la frontière du crépitation décidée par la FD est une définition approximative. Il est à noter que la différence d'IDW entre les deux types de crépitations est inférieure à 1 ms, une mesure précise des frontières des crépitations est donc nécessaire pour obtenir une classification

valide. Malheureusement, le crépitan est souvent distordu par le signal respiratoire, en particulier pour ceux de faible amplitude. Par conséquent, il devient plus difficile de calculer les paramètres IDW et 2CD de façon automatique. Selon l'expérimentation de Hoovers et Loudon ^[64], une discrimination basée sur la durée de la déflexion maximale pourrait présenter de meilleures performances que celle basée sur le début du crépitan.

Dans ce projet, nous choisissons la durée de la déflexion maximale (MDW) et la durée de la déflexion totale (TDW) comme paramètres de discrimination lors de la classification. La mesure automatique de ces paramètres a été proposée par Vannuccini et al. ^[57]. Elle se base sur le calcul de la première dérivée du signal en utilisant un filtre FIR de la famille *Savitzky-Golay* ^[57]. Cette famille de filtres est de type passe-bas, bien adaptée au lissage et à la différentiation, et dont les propriétés sont définies dans le domaine du temps au lieu du domaine de Fourier ^[57]. Ces filtres sont principalement caractérisés par le nombre de coefficients p , le degré du polynôme k et l'ordre de la dérivée d ^[57]. Par expérimentation, on a choisi les paramètres $p = 9$ et $k = 2$. La MDW est calculée par deux tangentes qui ont la pente maximale; tandis que la TDW est calculée par un seuil K égalant 0.1 fois l'amplitude maximale du crépitan.

En plus des deux paramètres temporels MDW et TDW, nous avons utilisé deux paramètres fréquentiels qui sont la fréquence du pic (PF) et la largeur de la bande gaussienne (GBW), définie par les fréquences correspondantes à l'atténuation de 3 dB de l'amplitude maximale. Le spectre du signal est calculé par la transformée de Fourier. La figure 3.6 représente un exemple de crépitan et les différents paramètres des domaines temps et fréquence.

3.4.2 Modèle de Prony

Le modèle Prony permet de décomposer, dans le domaine temporel, un signal $x(t)$ en une somme pondérée de sinusoides amorties :

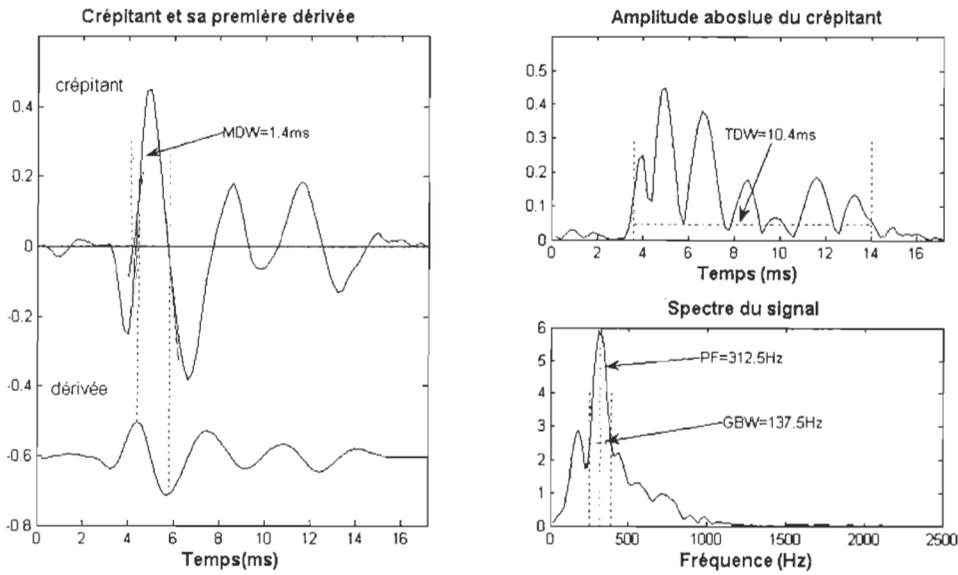


Figure 3.6: Caractérisation temps-fréquence d'un crépitan, où TDW=10.4ms, MDW=1.4ms, PF= 312.5Hz, GBW=137.5Hz

$$\hat{x}(t) = \sum_{k=1}^P A_k e^{a_k(t)} \cos(2\pi f_k t + \theta_k) \quad (3.15)$$

où p est l'ordre du modèle, A_k est le facteur de pondération, a_k est le coefficient d'atténuation, f_k et θ_k sont respectivement la fréquence et la phase de la $k^{\text{ième}}$ sinusoïde. De ce fait, le modèle Prony contient une série de sinusoïdes amorties, similaires à une moitié de crépitan.

La difficulté de cette technique consiste à déterminer les paramètres p , A_k , a_k , f_k et θ_k qui minimisent l'erreur quadratique :

$$\varepsilon = \sum_{n=0}^{N-1} |x(t) - \hat{x}(t)|^2 \quad (3.16)$$

Dans cette approche, le crépitan est divisé en deux parties afin de définir deux modèles. La première partie du crépitan est comprise entre l'instant de son commencement et l'instant où l'amplitude est maximale ; le restant du crépitan forme la deuxième

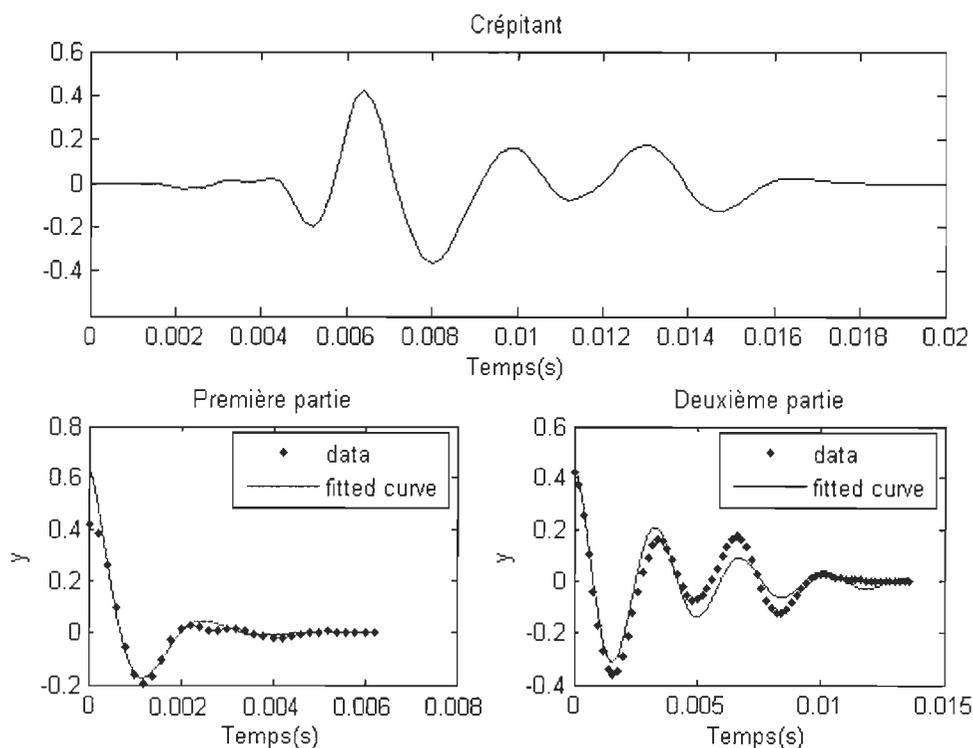


Figure 3.7: Modélisation par modèle de Prony

partie ^[66]. Pour modéliser le crépitant, la première partie du crépitant est inversée dans le temps. Après plusieurs tests sur l'ordre du modèle, la valeur de p est fixée à 1 parce qu'elle offre une erreur minimale ^[66]. Alors, nous avons quatre coefficients $\{A, a, f, \theta\}$ à ajuster dans chaque modèle de Prony. Dans notre expérimentation, les coefficients sont initialisés à $\{1, -150, 100, 0\}$, l'algorithme d'optimisation de ces coefficients provient de la boîte à outils "Curve Fitting Toolbox" de Matlab.

Après modélisation, on a 8 coefficients représentant un seul crépitant. Les coefficients des modèles de Prony pourraient être utilisés pour discriminer le crépitant en gros crépitant ou en crépitant fin. La figure 3.7 représente un crépitant réel et le résultat de la modélisation par le modèle de Prony.

Le problème rencontré avec cette méthode réside dans son instabilité. À cause de la présence du bruit et de la distorsion introduite par le filtre de séparation, la fonction de Prony n'est pas toujours un modèle idéal pour estimer le crépitant. Si le signal crépitant

est trop loin d'une forme sinusoïde amortie, le résultat de la modélisation ne peut pas présenter une bonne caractérisation du crépitan et aboutit à une mauvaise classification.

3.4.3 Transformée en ondelettes continue

Nous rappelons que la transformée en ondelettes continue d'un signal $x(t)$ est une fonction à deux variables définie par :

$$w(\tau, s) = \frac{1}{s} \int x(t) \psi\left(\frac{t - \tau}{s}\right) dt \quad (3.17)$$

où $\psi(t)$ est l'ondelette mère, s est le paramètre échelle et τ est le paramètre de translation de la fonction ondelette le long de l'axe du temps.

L'ondelette mère $\psi(t)$ est une petite onde amortie, localisée en temps et en fréquence. Par exemple, la figure 3.8 représente la composante réelle de l'ondelette de Morlet définie par : $\psi(t) = e^{-\alpha t^2/2} e^{j\omega_0 t}$, où $\alpha = 1$ et $\omega_0 = 5$ déterminent respectivement l'enveloppe et l'oscillation de l'ondelette.

Du et al. ^[67] ont proposé une méthode de classification de crépitants basée sur la transformée en ondelettes continue. En fait, ils partent de l'hypothèse que la transformée en ondelettes continue peut être considérée comme une "inter-corrélation" entre $x(t)$ et $\frac{1}{s}\psi(\frac{t}{s})$, par conséquent, si un modèle de crépitan est utilisé comme ondelette mère, la transformée en ondelettes devient une *auto-corrélation* ^[67]. Avec la variation du paramètre échelle s , l'ondelette $\frac{1}{s}\psi(\frac{t}{s})$ qui possède une fréquence d'ondulation plus proche de celle du crépitan fournira une réponse optimale au niveau de la fonction d'intercorrélacion et permettra ainsi de le localiser. Nous savons que le crépitan fin est caractérisé par une fréquence plus élevée que celle du gros crépitan. L'idée principale derrière cette approche consiste à utiliser le paramètre d'échelle s comme seuil pour discriminer les crépitants fins des gros crépitants.

Normalement un crépitan possède entre 4 et 10 déflexions. La durée d'une déflexion est plus courte au début du crépitan et devient plus longue ensuite. L'amplitude du signal

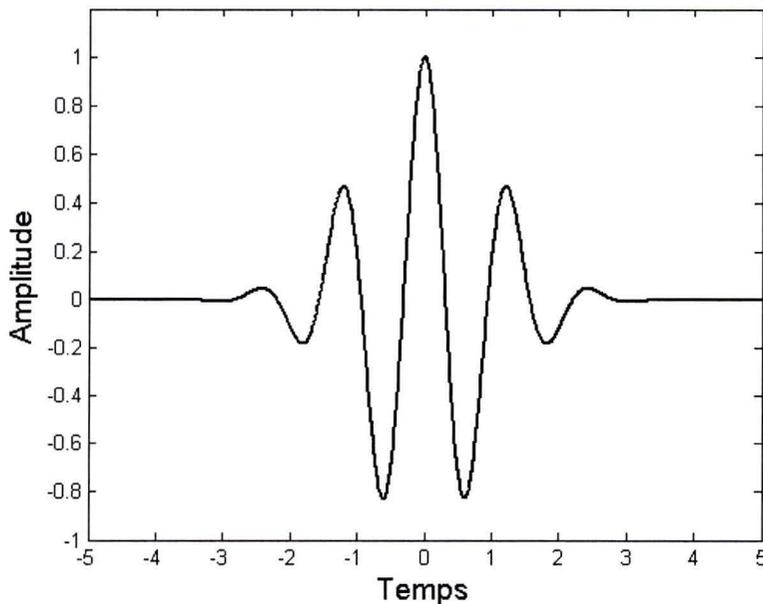


Figure 3.8: La composante réelle de l'ondelette de Morlet

augmente au début et diminue après que l'amplitude maximale est atteinte. Une fonction de simulation d'un crépitan est définie par ^[67] :

$$g(t) = \sin(2\pi * f_0 * t^{a_1}) * t^{a_2} * e^{t/a_3} \quad (3.18)$$

où $a_1 = 0.5$; $a_2 = 1.49$; $a_3 = 0.78$ et $f_0 = 2.0$.

La figure 3.9 représente une forme d'onde obtenue pour $s = 3$. C'est une forme qui ressemble étroitement à celle d'un crépitan et qui pourrait être utilisée comme ondelette mère.

Pour un crépitan donné à classifier, on calcule la transformée en ondelettes avant de déterminer l'échelle optimale pour laquelle l'énergie de la fonction $w(\tau, s)$ est maximale. Afin de discriminer les crépitans, on compare leurs échelles optimales à un seuil fixé empiriquement à 3.1. Les crépitans qui ont une échelle optimale inférieure à 3.1 sont considérés comme des crépitans fins, les autres sont considérés comme étant de gros crépitans.

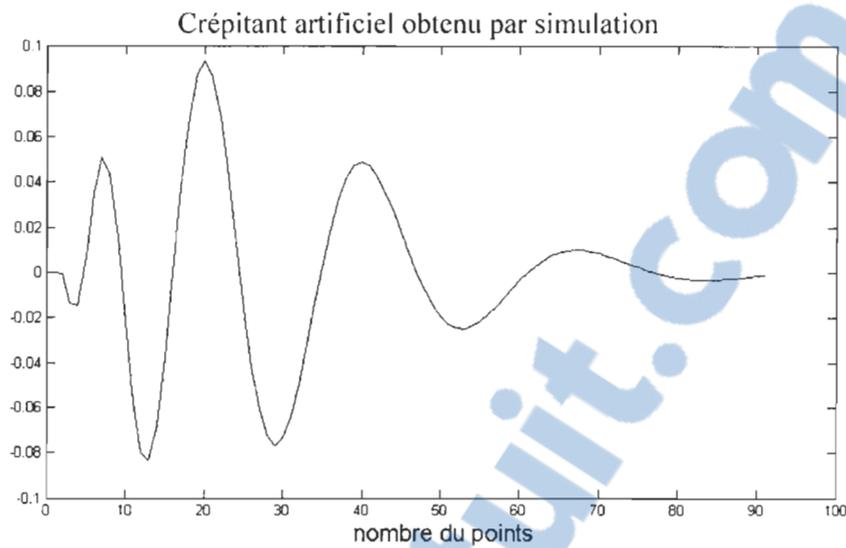


Figure 3.9: Crépitant simulé $\psi(\frac{t}{s})$ obtenu par la fonction $\psi(t) = g(t)$ et $s = 3$.

La figure 3.10 présente l'analyse par ondelettes, sur plusieurs échelles, d'un crépitant réel en utilisant la fonction $g(t)$ comme ondelette mère. L'énergie maximale est obtenue quand l'échelle est égale à 2,4. Alors, on pourrait déduire qu'il s'agit d'un crépitant fin.

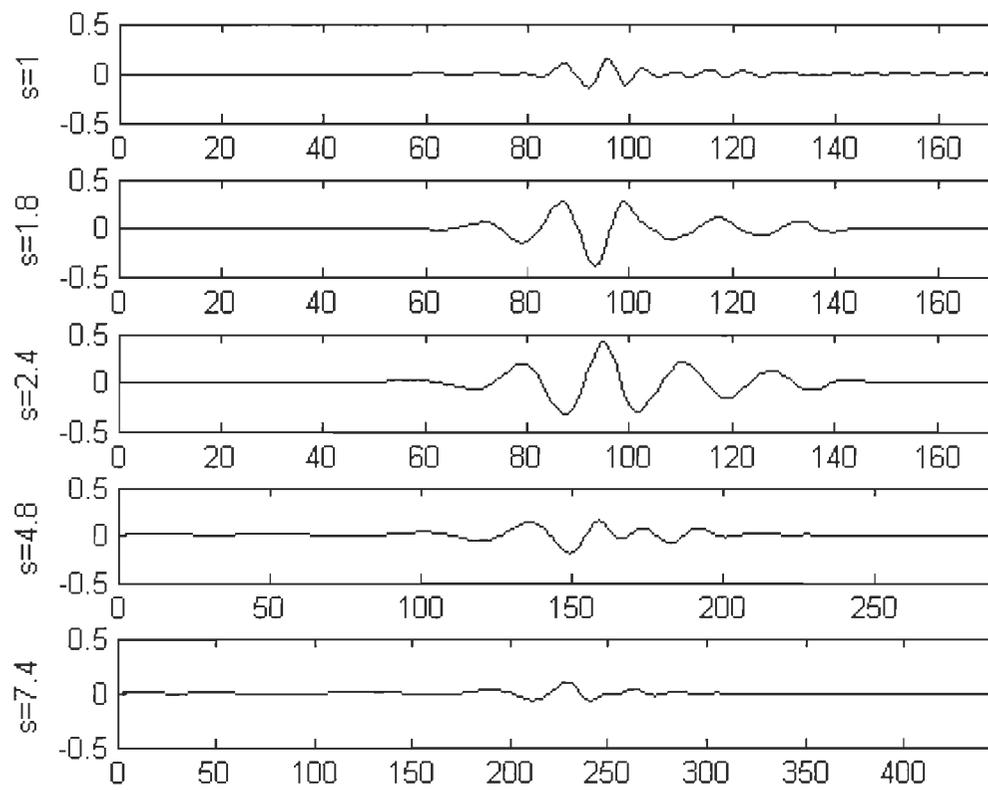


Figure 3.10: Analyse par ondelettes, sur plusieurs échelles, d'un crépitant réel en utilisant la fonction $g(t)$ comme ondelette mère.

CHAPITRE IV

DÉTECTION DES SIBILANTS

Ce chapitre traite les méthodes de détection des sibilants. Trois de ces méthodes sont basées sur l'analyse du spectrogramme, la quatrième utilise la représentation par "Chirplets". Ces techniques peuvent être utilisées comme une première étape (prétraitement) dans un système de reconnaissance automatiques des sibilants.

4.1 Principe

Les sibilants sont des sons pulmonaires adventices continus, qui se superposent aux sons respiratoires normaux [2]. Ils sont souvent liés à l'obstruction des voies aériennes des poumons et associés généralement à une limitation du débit. Selon la définition de l'ATS (*American Thoracic Society*), la durée minimale d'un sibilant est de 250 ms, et sa fréquence dominante est supérieure à 400 Hz (voir le tableau 1.1). Les récentes recherches démontrent que 250 ms semble trop long pour une durée minimum du sibilant. Le groupe CORSA (*Computerized Respiratory Sound Analysis*) définit la durée minimale acceptable d'un sibilant comme étant de 80 à 100 ms et la fréquence dominante minimale de 100 Hz [22]. On appelle le sibilant contenant un seul harmonique "monophonique", alors que celui ayant plusieurs harmoniques est appelé "polyphonique".

Le sibilant est étroitement lié à l'asthme mais d'autres causes existent, telles que les tumeurs, l'inhalation de corps étrangers, l'embolie pulmonaire, défaillance cardiaque congestive. Plusieurs travaux ont été menés pour étudier l'effet de la sévérité de l'asthme sur la fréquence, la durée, le nombre et l'intensité des sibilants. À titre d'exemple, Baugh-

man et Loudon ^[68] trouvent que les meilleurs marqueurs étaient la proportion du cycle respiratoire occupé par le sibilant (T_s/T_{tot}), et la fréquence du pic dominant du sibilant. En effet, en réduisant l'obstruction à l'aide d'un bronchodilatateur, ils ont montré que, le rapport T_s/T_{tot} et la fréquence du pic dominant sont tous les deux réduits ^[50].

4.2 Calcul du spectrogramme

La plupart des méthodes utilisées dans ce chapitre se basent sur l'analyse du spectrogramme, calculé à l'aide de la transformée de Fourier à court terme (STFT).

Le signal respiratoire à temps-discret $x[n]$ est multiplié par une fenêtre glissante $w[n]$ de L échantillons, qui avance avec un pas de K échantillons. Pour chaque position n de cette fenêtre, la transformée de Fourier du segment obtenu, est calculée selon l'équation suivante :

$$X[n, k] = \sum_{m=0}^{M-1} x[nL + m]w[m]e^{-j\frac{2\pi mk}{M}} \quad (4.1)$$

avec $n = 0, 1, \dots, N - 1$, $k = 0, 1, \dots, M - 1$, où n est l'indice du temps, k indice des fréquences, N indice du dernier segment et M est le nombre de fréquences discrètes. En pratique, on prend $M = L$. Plusieurs fenêtres d'apodisation sont proposées dans la littérature, mais on utilise souvent les fenêtres de Hamming ou de Hanning. Le taux de chevauchement (entrelacement) est défini par K/L . Pour une fréquence d'échantillonnage F_e , les résolutions temporelle et fréquentielle du spectrogramme sont respectivement données par $\Delta t = K/F_e$ et $\Delta f = F_e/(2M)$.

Le spectrogramme $SP_x[n, k]$ du signal est une représentation qui illustre la distribution temps-fréquence de l'énergie du signal $SP_x[n, k] = |X[n, k]|^2$. On utilise aussi le module des transformées de Fourier pour représenter le spectrogramme $SP_x[n, k] = |X[n, k]|$. Le spectrogramme d'un signal est une matrice à deux dimensions (2D) qui peut être traitée comme une image.

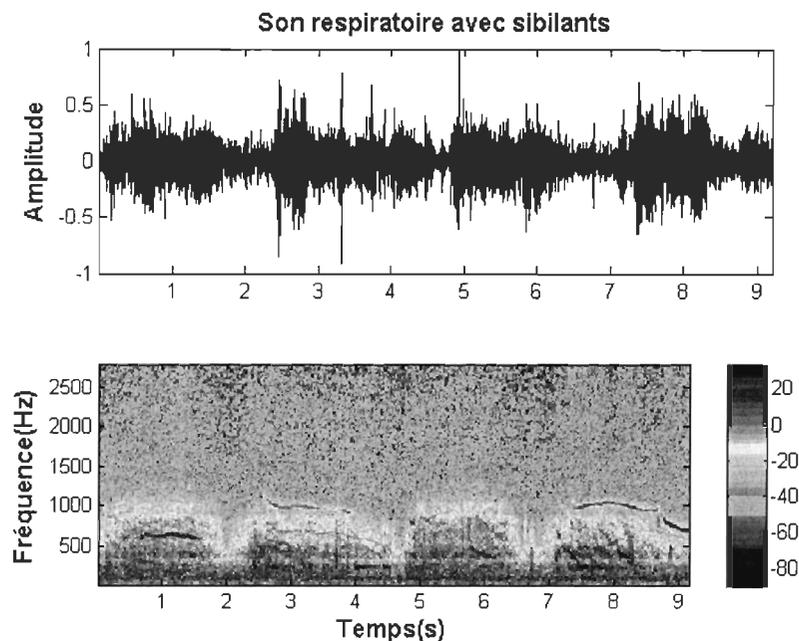


Figure 4.1: Exemple de spectrogramme d'un son respiratoire avec sibilants calculé à l'aide de la STFT, où $F_e = 8000$ Hz, $L = M = 512$ et le taux de chevauchement est de 75 %

La présence d'un sibilant se traduit par une rayure dans le spectrogramme correspondant à un pic assez fort ayant une certaine durée. Dans ce type d'approche, la détection du sibilant est réalisée en utilisant un seuil de discrimination à partir duquel les pics sont acceptés ou rejetés. Alors, la définition du seuil est l'essence dans la détection du sibilant.

Dans ce chapitre, quatre méthodes de détection des sibilants sont discutées : le Laplacien d'une gaussienne ^[42], la modélisation auditive ^[43], la détection d'épisodes de sibilants ^[44] et la représentation par "Chirplets" ^[45].

4.3 Détection des sibilants par le Laplacien d'une gaussienne

En considérant le spectrogramme comme une image 2D, Riella et al. ^[42] proposent l'utilisation des filtres bidimensionnels en vue de détecter les sibilants. Ils utilisent essentiellement un masque LoG (*Laplacian of Gaussian*) dans le but d'augmenter le contraste de l'image.

4.3.1 Laplacien d'une gaussienne (LoG)

Le Laplacien d'une gaussienne (LoG) est un opérateur obtenu par combinaison de l'opérateur Laplacien et d'un filtre de lissage gaussien ; il est souvent utilisé dans la détection des contours d'une image. Les contours sont détectés aux passages par zéro du LoG, où le plus souvent l'intensité de l'image change rapidement.

Pour une image $I(x, y)$, le Laplacien $L(x, y)$ est une mesure isotrope de la deuxième dérivée partielle.

$$L(x, y) = \nabla^2 I(x, y) = \frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2} \quad (4.2)$$

La dérivée étant trop sensible au bruit, un filtre de lissage $G(x, y)$ est utilisé pour réduire l'effet du bruit dans l'image .

$$\nabla^2(I * G) = (\nabla^2 I) * G = I * (\nabla^2 G) \quad (4.3)$$

Les opérations de filtrage et de dérivation se font donc en une seule étape. Le plus souvent, on utilise un filtre de lissage gaussien de variance σ^2 :

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (4.4)$$

L'opérateur composé $LoG(x, y)$ est défini par :

$$LoG(x, y) = -\frac{1}{\pi\sigma^4} \left[1 - \frac{x^2 + y^2}{2\sigma^2} \right] e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (4.5)$$

L'écart type σ détermine la précision de détection des arêtes. Avec un σ élevé, seules les arêtes dominantes sont détectées, alors qu'avec un σ faible, toutes les arêtes sont détectées avec précision.

Pour traiter une image numérique, l'opérateur LoG peut être représenté approximati-

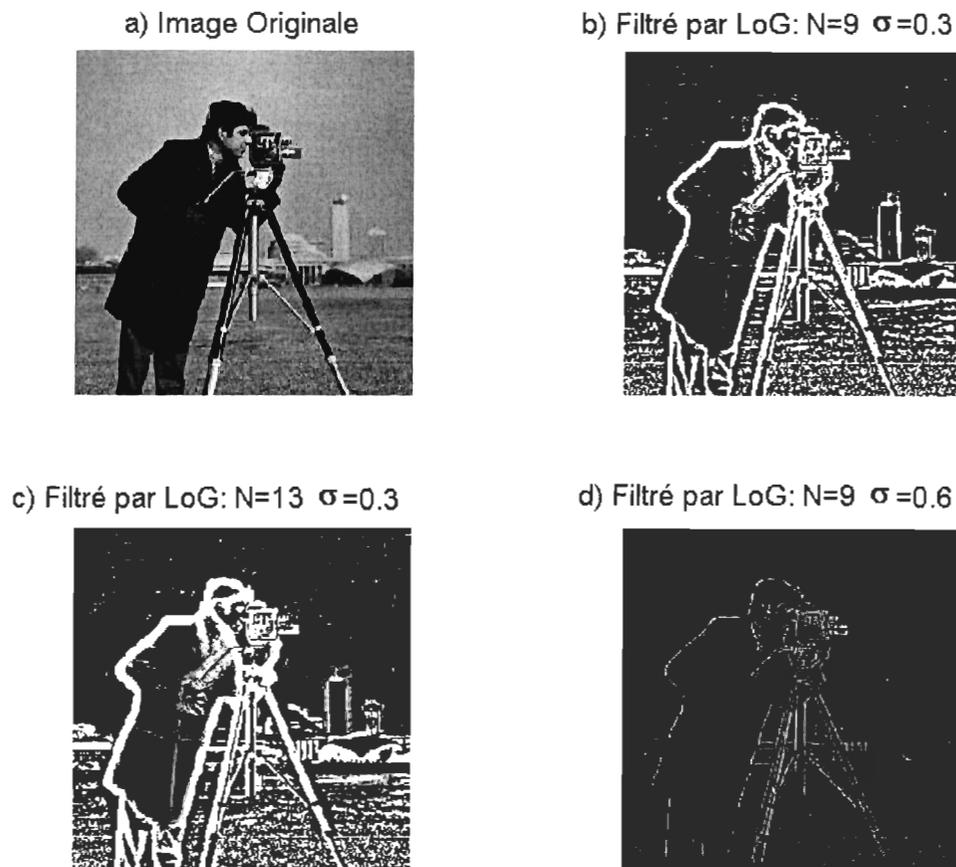


Figure 4.2: Filtrage d'images par le Laplacien d'une gaussienne (LoG)

vement par une matrice carrée de dimension N . La figure 4.2 donne un exemple d'images filtrées par le Laplacien d'une gaussienne (LoG). La dimension N détermine la largeur des contours des images filtrées, alors que σ détermine du niveau de lissage (voir la figure 4.2).

4.3.2 Détection des sibilants par filtre LoG

La détection des sibilants dans un son respiratoire se résume en trois étapes :

- Découpage du signal en plusieurs segments entrelacés, puis application de la transformée de Fourier à court terme (STFT) à chacun des segments. Le spectrogramme représente le module de ces transformées dans le plan temps-fréquence (figure 4.3-a).
- Convolution du spectrogramme obtenu avec un masque LoG, afin de rehausser les

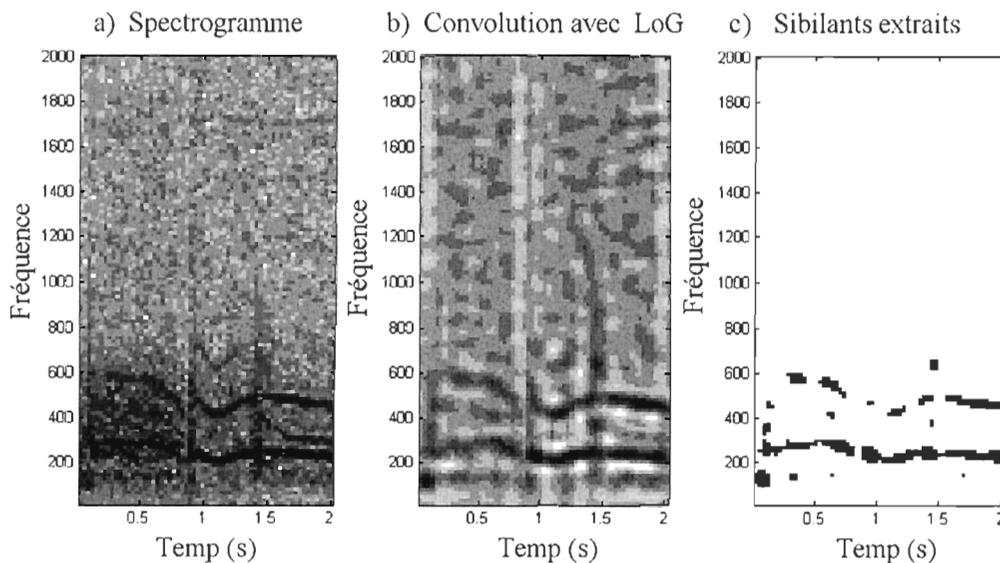


Figure 4.3: Détection des sibilants en utilisant le Laplacien d'une gaussienne (LoG). Les paramètres utilisés sont : $F_e = 8000$ Hz, $L = M = 512$, taux de chevauchement de 50%, $N = 9$, $\sigma = 2.5$ et $p = 0.4$

contours des sibilants et lisser les bruits (figure 4.3-b). Nous avons utilisé un masque $[9 \times 9]$ correspondant à $N = 9$ et $\sigma = 2.5$.

- Application d'un seuil universel proportionnel à l'amplitude maximum du spectrogramme filtré, pour ne garder que les sibilants (figure 4.3-c).

$$Seuil = p * (max(SP_x[n, k])) \quad (4.6)$$

où p est le coefficient de proportionnalité.

En réalité, la détection d'un sibilant n'est pas vraiment un problème de détection de contours. En effet, le filtrage LoG accentue les valeurs du spectrogramme (intensité) au milieu des contours du sibilant. Le seuillage permet d'extraire le sibilant mais pas ses contours, qui peuvent être détectés aux passages par zéro du LoG.

L'exemple de la figure 4.3 correspond à un cycle respiratoire contenant trois sibilants : un sibilant polyphonique de fréquences 300 Hz et 600 Hz est présent à l'inspiration, un sibilant polyphonique de fréquences 220 Hz, 440 Hz et 640 Hz est présent à l'expiration et

un sibilant faible en amplitude et d'une fréquence d'environ 300 Hz, apparaissant à la fin de l'expiration. La figure 4.3-c montre que cette technique extrait la plupart des contours des sibilants forts en amplitude, mais introduit un certain nombre de fausses détections.

L'avantage de cette méthode réside dans la simplicité de son algorithme. Le masque LoG est une matrice constante et le seuil de discrimination prend une valeur universelle. Le calcul principal se réduit à une convolution 2D. Ses inconvénients sont l'inflexibilité de son algorithme et son incapacité de détecter les sibilants faibles. La performance de détection est directement influencée par le paramètre p , qui dépend de la durée et de la qualité du signal (niveau du sibilant dans le son respiratoire). Il est alors difficile de choisir une valeur de p fixe pour les sons respiratoires. D'ailleurs, cette méthode est plus susceptible d'ignorer les sibilants de faible amplitude à cause de la valeur universelle du seuil.

4.4 Détection des sibilants par modélisation auditive

4.4.1 Système auditif

Le système périphérique auditif d'un humain peut être modélisé par une batterie de filtres passe-bande entrelacés, appelés filtres auditifs ^[43]. Un son tonal est audible quand son niveau de pression sonore SPL (*sound pressure level*) est au-dessus d'un seuil de l'ouïe et persiste plus de 10 ms ^[43]. Quand un auditeur essaie de détecter un son tonal masqué par un bruit, il est supposé utiliser le filtre ayant une fréquence centrale plus proche de celle du signal tonal. Comme le bruit présente un effet de masquage, il augmente le seuil SPL auquel le son devient audible, mais seulement les composantes fréquentielles du bruit dans la bande du filtre auditif correspondant ^[43].

La bande-passante des filtres auditifs dépend de la fréquence. Glasberg et Moore ^[69] proposent une mesure de la largeur de bande de ces filtres, appelée *Equivalent Rectangular*

Bandwidth (ERB). Ces largeurs de bandes ERB sont définies par :

$$ERB_f = 0.108f + 24.7 \quad (4.7)$$

où f est la fréquence centrale du filtre en Hz.

4.4.2 Description de l'algorithme (FDDT)

Qiu et al. ^[43] propose une nouvelle méthode de détection des sibilants, appelée FDDT (*Frequency-and duration-dependant threshold*). Elle se base sur le principe qu'un sibilant audible peut être identifié par la présence d'un pic dans le spectrogramme, dépassant un certain seuil. La définition de ce seuil se base sur les résultats expérimentaux obtenus par Reed et Bilger ^[70], qui trouvent que le seuil de masquage dépend de la fréquence et du niveau de bruit. Autrement dit, pour détecter un son tonal, son rapport signal-sur-bruit SNR devrait dépasser un certain seuil qui dépend de la fréquence centrale du signal. Le SNR varie de 8 dB pour 250 Hz à 14 dB pour 4 kHz ^[70]. Qiu et al. ^[43] définissent un seuil dépendant de la fréquence et de l'ERB :

$$Seuil_1 = 9.5 + 3.48\log_{10}(0.01f) - 10\log_{10}(ERB_f) \quad (4.8)$$

où la fréquence f et l' ERB_f sont en Hz.

Il est à noter que cette définition ne correspond pas aux résultats de Reed et Bilger ^[70]. Avec cette équation, le seuil est de -6.25 dB pour 250 Hz et de -21.30 dB pour 4 kHz. Pour notre projet, nous avons défini un nouveau seuil :

$$Seuil_2 = 5.5 + 12.87\log_{10}(0.0652f) - 10\log_{10}(ERB_f) \quad (4.9)$$

Ce seuil varie de 4 dB pour 250 Hz à 10 dB pour 4 kHz. Comparé au résultat de Reed et Bilger (8 dB pour 250 Hz à 14 dB pour 4 kHz), nous avons diminué le seuil de 4 dB pour avoir une détection plus précise. Dans le cas de la détection des sibilants, les pics du

spectrogramme sont considérés comme le signal et le reste comme du bruit. Le SNR est calculé dans la bande de fréquences définie par l'ERB. Si le SNR du pic dépasse le seuil, il est détecté comme un sibilant.

Illustrée à la figure 4.4, cette technique est décrite en détails par l'algorithme ci-dessous :

1. Construire une matrice temps-fréquence A , contenant l'énergie du signal ($|X[n, k]|^2$). Les paramètres de la STFT sont : fréquence d'échantillonnage de 8000 Hz, fenêtre de Hanning, et une largeur de la fenêtre de 512 échantillons avec entrelacement de 50%. Initialiser la matrice de masquage $B = 0$, qui a la même taille que A . La résolution temporelle du spectrogramme est $\Delta t = 512/8000 = 32$ ms.
2. Trouver les pics en parcourant les colonnes de la matrice A . Dans la colonne j , si $A(i, j) > A(i + 1, j)$ et $A(i, j) > A(i - 1, j)$, $A(i, j)$ est un pic.
3. Calculer le rapport signal sur bruit SNR. L'énergie du signal est égale à celle du pic et l'énergie du bruit est égale à celle du son respiratoire dans l'ERB correspondant. Si le SNR excède le seuil, marquer $B(i, j) = 1$.
4. Étiqueter les pics connectés par l'algorithme CCL (*Connected Component labeling*)^[71]. Le principe de cet algorithme consiste à donner une étiquette (*label*) aux pics appartenant au même sibilant.
5. En se basant sur la durée des signaux ayant la même étiquette (*label*), utiliser la durée D et le SNR pour supprimer les étiquettes de la matrice B , qui correspondent à des signaux de courte durée et n'excédant pas certains seuils de discrimination. Si la durée $D = \Delta t$ et $SNR < Seuil_2 + 4.8$ ou $D = 2\Delta t$ et $SNR < Seuil_2 + 3$, alors enlever ce sibilant de la matrice B . Garder les sibilants ayant une durée $D \geq 3\Delta t$.
6. Les composantes détectées dans l'étape 5 peuvent appartenir au même sibilant. Pour les connecter, nous devons re-vérifier les pics détectés dans l'étape 2. Si le pic avoisine le sibilant détecté, étiqueter le par la même étiquette (*label*). La définition des voisins est semblable à celle de l'algorithme CCL.

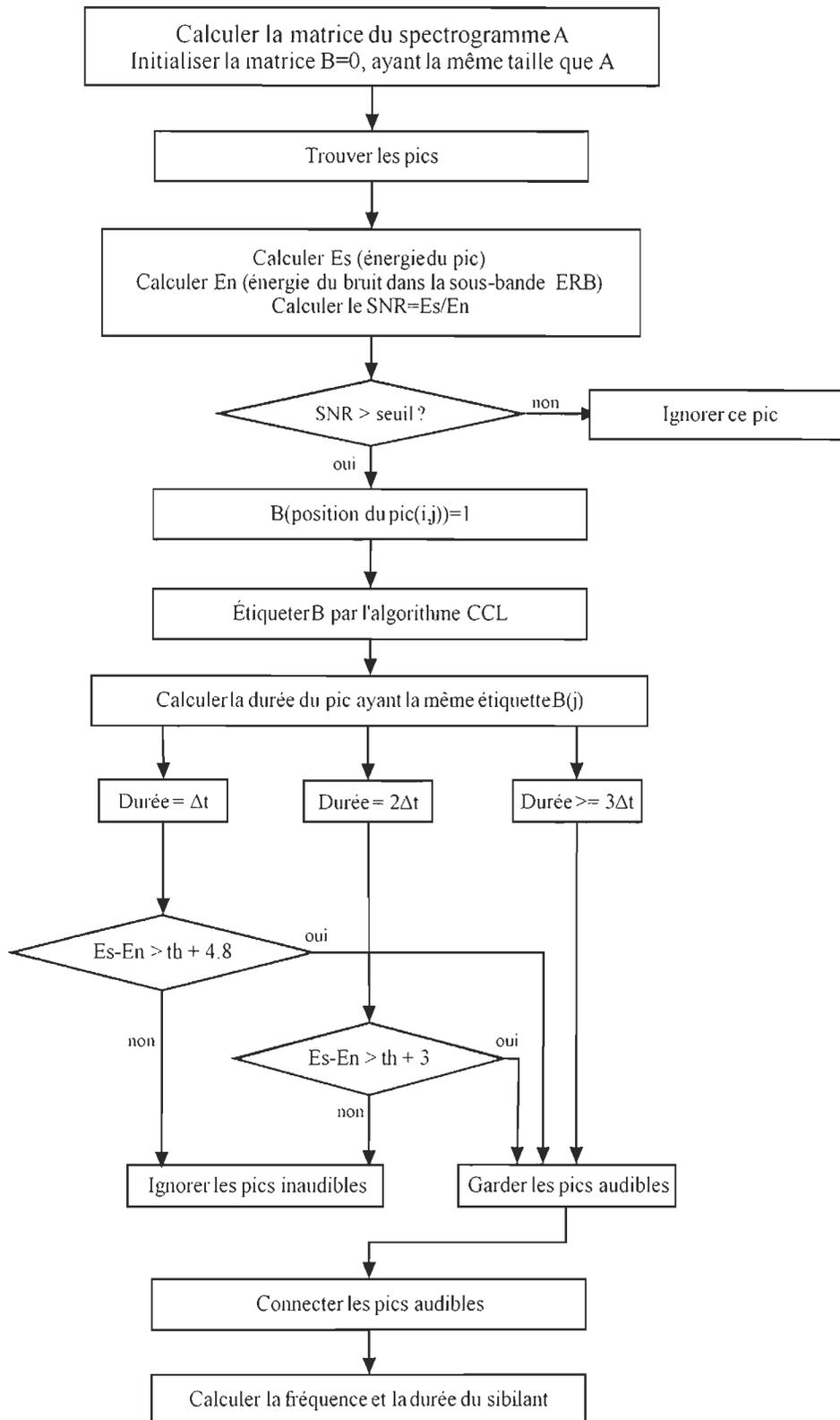


Figure 4.4: Schéma de la détection des sibilants par modélisation auditive (FDDT).

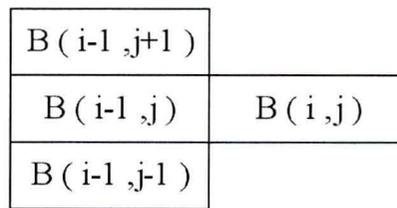


Figure 4.5: Définition des voisins pour l'algorithme CCL

7. Calculer la durée et la fréquence moyenne du sibilant détecté. On pourrait facilement avoir la proportion du cycle respiratoire occupé par le sibilant (T_s/T_{tot}).

L'algorithme CCL ^[71] regroupe les pics connectés dans le spectrogramme. Il parcourt la matrice B , colonne par colonne, et assigne une valeur du groupement à chaque élément non nul (voir la figure 4.5).

1. Pour un élément $B(i, j) = 1$, on examine ses trois voisins : $B(i-1, j-1)$, $B(i-1, j)$ et $B(i-1, j+1)$. Seulement un voisin au maximum pourrait être non nul.
2. Si $B(i, j) = 1$ et tous les voisins du $B(i, j)$ valent 0, alors un nouveau sibilant est détecté. Par conséquent, un nouveau label est assigné au $B(i, j)$.
3. Si $B(i, j) = 1$ et ses voisins ont un label non nul, alors ils viennent du même sibilant. Un même label est assigné au $B(i, j)$.

Un exemple de détection de sibilants par modélisation auditive (FDDT) est représenté à la figure 4.6-a, qui reprend le même signal que l'exemple du filtre LoG. Ce signal est caractérisé par : un sibilant polyphonique (I) dans la phase d'inspiration (aux fréquences 300 Hz et 600 Hz environ), un sibilant polyphonique (II) dans la phase d'expiration (aux fréquences 220 Hz, 440 Hz et 640 Hz environ) et un sibilant monophonique (III) qui apparaît à la fin d'expiration (à la fréquence 300 Hz environ). Cette méthode détecte 9 sibilants, numéroté de 1 à 9 dans la figure 4.6-b. Les sibilants détectés 1 et 2 proviennent du sibilant I ; les sibilants détectés 4, 5, 6, 7 et 9 proviennent du sibilant II ; les sibilants 3 et 8 sont des fausses détections ; et enfin le sibilant III a été ignoré (ou perdu) dans cette expérimentation. La méthode FDDT définit un seuil dépendant de la fréquence,

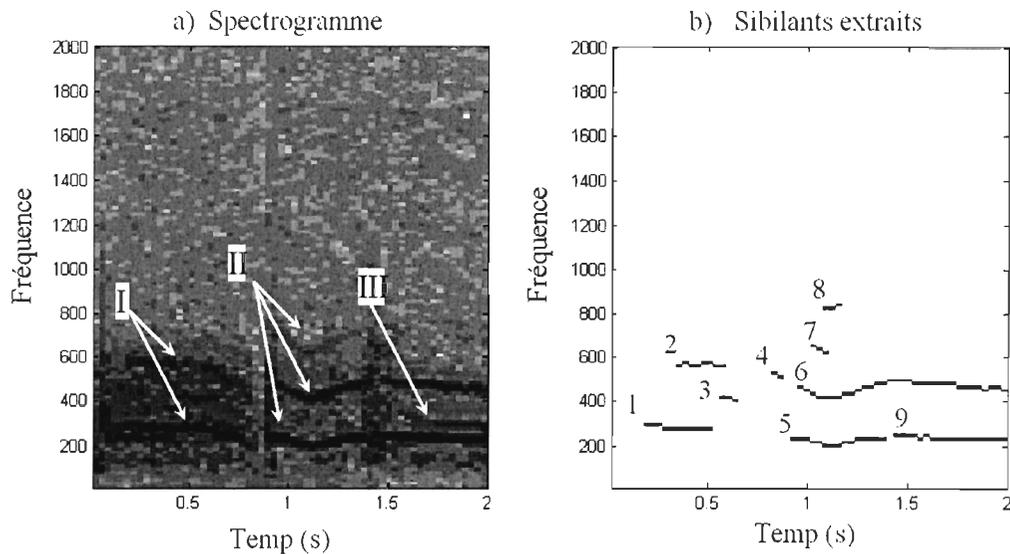


Figure 4.6: Exemple de détection des sibilants par modélisation auditive (FDDT). Les paramètres utilisés sont : $F_e = 8000$ Hz, $L = M = 512$ et le taux de chevauchement de 50 %.

elle extrait la plupart des sibilants avec moins d'erreur. Mais elle présente l'inconvénient d'ignorer les sibilants faibles (par exemple, le troisième harmonique du sibilant II et le sibilant III de la figure 4.6-a).

4.5 Détecteur d'épisodes des sibilants (WED)

Taplidou et al. ^[44] proposent une méthode améliorée pour détecter les épisodes des sibilants, appelée WED (*Wheezing episodes detector*). Semblable à la méthode FDDT de la section 4.4, elle se base sur la détection des pics dans le spectrogramme du signal. La différence réside dans la définition d'une bande de fréquence, susceptible de contenir les sibilants, et son découpage en 6 sous-bandes. Dans chaque sous-bande, le seuil de discrimination est calculé selon la moyenne et l'écart-type de l'amplitude du signal. D'autre part, ils définissent la continuité temporelle et fréquentielle du spectre, qui supprime efficacement les fausses détections et connecte les fragments disjointes d'un sibilant.

Illustré à la figure 4.7, cet algorithme est décrit comme suit :

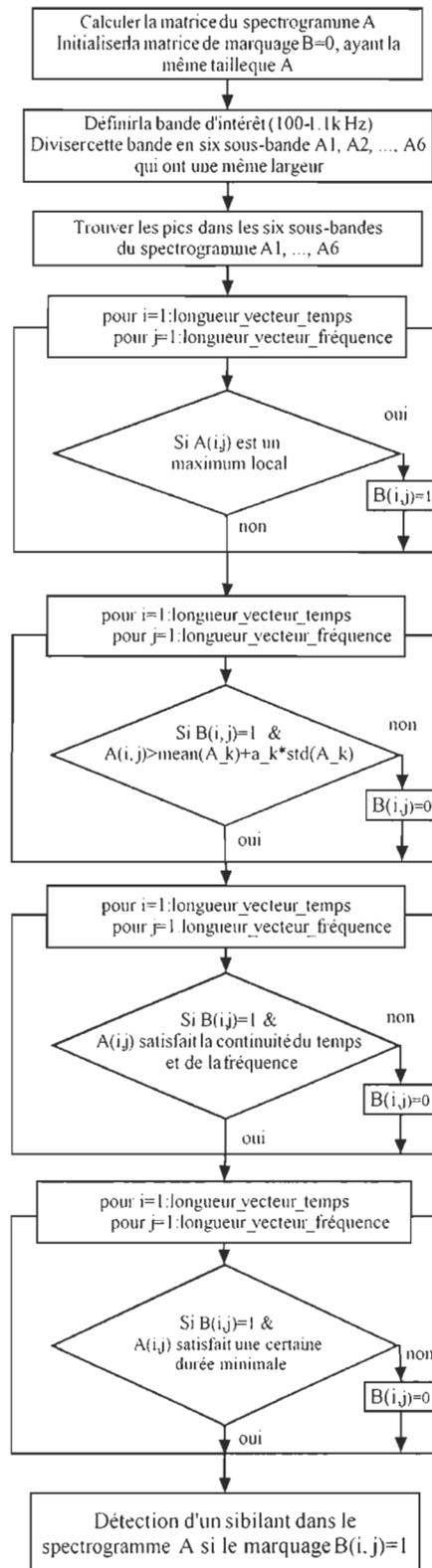


Figure 4.7: Schéma du détecteur d'épisodes de sibilants (WED).

1. Construire la matrice spectrogramme A , qui contient l'amplitude de spectre du signal en dB. Les paramètres de la STFT sont : taux d'échantillonnage de 8000 Hz, fenêtre de type Hanning de longueur $L=256$, longueur de la FFT, $M=2048$, et un entrelacement de 50%. Initialiser la matrice du masque $B=0$, de même taille que la matrice A .
2. En se limitant à la bande d'intérêt (100 – 1100 Hz), trouver les maximums locaux le long de chaque colonne de la matrice A . Dans une colonne j , un point $A(i, j)$ est considéré comme un maximum local s'il vérifie la condition : $A(i, j) > A(i + 1, j)$ et $A(i, j) > A(i - 1, j)$. Marquer les maximums locaux par des "1" dans la matrice B .
3. Diviser la bande d'intérêt en six sous-bandes. Dans chaque sous-bande k , calculer le seuil $Th_k = mean(A) + a_k * std(A)$, enlever les maximums locaux qui sont inférieures au seuil correspondant, en forçant leur marque à "0" dans la matrice B .
4. Regrouper les pics selon la continuité temporelle et fréquentielle. Si l'écart temporel Δt_{WED} entre deux pics est inférieur à 21 ms et l'écart fréquentiel Δf_{WED} est inférieur à 20 Hz, les deux pics appartiennent à un même sibilant.
5. Calculer la durée du sibilant détecté Δt_{sib} . Si la durée est inférieure à 150 ms, enlever ce sibilant de la matrice B .

Comme pour la deuxième méthode (section 4.4), le principe de cette méthode est basé sur l'utilisation d'un seuillage dans le spectrogramme dépendant de la fréquence du signal. Mais, elle ajoute de nouvelles restrictions pour vérifier la continuité temporelle et fréquentielle du sibilant dans le but d'assurer la validité de la détection. En utilisant le même signal que précédemment (figure 4.8-a), la méthode (WED) détecte 7 sibilants, numérotés de 1 à 7 dans la figure 4.8-b. Les sibilants détectés 1, 2 et 3 proviennent du sibilant I; les sibilants 4, 5 et 6 proviennent du sibilant II; le sibilant détecté 7 provient du sibilant III et enfin le troisième harmonique du sibilant II est perdu.

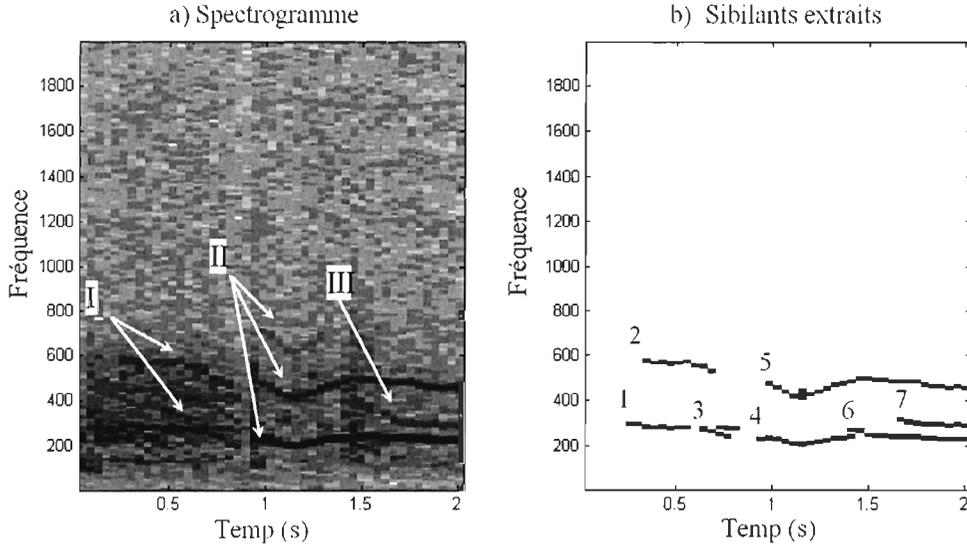


Figure 4.8: Résultat de la détection par le détecteur d'épisodes de sibilants (WED). Les paramètres utilisés sont : $F_e = 6000$ Hz, $L = 256$, $M = 2048$, le taux de chevauchement est de 50 %, $\Delta f_{WED} = 25$ Hz, $\Delta t_{WED} = 21$ ms, $\Delta t_{sib} = 150$ ms, $a_k = [1.5, 1.5, 1.5, 1.5, 1.6, 1.6]$.

4.6 Représentation des sibilants par "Chirplets"

4.6.1 Transformée par "Chirplets"

La transformée par *Chirplets* a été proposée, pour la première fois, par Mann et Haykin [72,73] pour analyser les signaux radars. C'est une technique d'analyse temps-fréquence qui consiste à décomposer le signal en utilisant une base de fonctions, appelées "Chirplets". Dans le cas de la transformée par ondelettes (sections 2.4.1 et 3.4.3), la base d'ondelettes est dérivée à partir d'une ondelette mère par les opérations de dilation et de translation. De même, la base de chirplets d'une transformée gaussienne par chirplets est dérivée d'une simple fonction gaussienne $g(t) = \pi^{-1/4} \exp(-t^2/2)$ par quatre opérations : changement d'échelle, "chirping", décalage en temps et décalage en fréquence [74]. Ce qui conduit à une famille d'ondes à quatre paramètres d'ajustement [75] :

$$g_{t_c, \omega_c, c, \Delta_t}(t) = \frac{1}{\sqrt{\sqrt{\pi} \Delta_t}} e^{-\frac{1}{2} \left(\frac{t-t_c}{\Delta_t} \right)^2} e^{j[c(t-t_c) + \omega_c(t-t_c)]} \quad (4.10)$$

où $j = \sqrt{-1}$, $t_c \in \mathbb{R}$ est la localisation dans le temps, $\omega_c \in \mathbb{R}$ la localisation en fréquence, Δ_t la durée (ou étendue, et c le taux de modulation du chirp qui caractérise le rapidité du changement de la fréquence ^[74]. Par conséquent, la transformée par chirplets d'un signal $x(t)$ est définie par le produit scalaire du signal par la chirplet $g_{t_c, \omega_c, c, \Delta_t}(t)$:

$$a_{t_c, \omega_c, c, \Delta_t} = \int_{-\infty}^{\infty} x(t) g_{t_c, \omega_c, c, \Delta_t}^*(t) dt \quad (4.11)$$

où $(*)$ représente le complexe conjugué. Les coefficients $a_{t_c, \omega_c, c, \Delta_t}$ représentent le contenu énergétique du signal dans les régions temps-fréquence spécifiées par les chirplets ^[74].

Pour alléger la notation, l'ensemble des paramètres du chirplet est décrit par $\beta = \{t_c, \omega_c, c, \Delta_t\}$. Un signal donné, $x(t)$, peut être reconstruit par une combinaison linéaire des chirplets.

$$\begin{aligned} x(t) &= \sum_{n=1}^P a_{\beta_n} g_{\beta_n}(t) + R^{P+1}x(t) \\ &= f_P(t) + R^{P+1}x(t) \end{aligned} \quad (4.12)$$

où β_n représente l'ensemble des paramètres du $n^{\text{ième}}$ chirplet, $R^{P+1}x(t)$ dénote le résidu et $f_P(t)$ est définie comme étant l'approximation d'ordre P du signal.

L'estimation optimale des a_{β_n} et β_n correspondant à la décomposition d'un signal $x(t)$ dans la base de fonctions g_{β_n} est un problème très complexe ^[74]. On utilise souvent deux algorithmes : l'estimation du maximum de vraisemblance MLE (*Maximum Likelihood Estimation*) et l'estimation-maximisation EM (*Expectation-Maximization*). Pour plus de détails, nous invitons le lecteur à consulter d'autres références ^[45, 74].

Pour implanter cette transformée, nous avons utilisé "DiscreteTFDs" ^[76], une boîte à outils Matlab du domaine public. La transformée par *chirplets* a été appliquée, entre autres, pour analyser des sifflements de baleines ^[45].

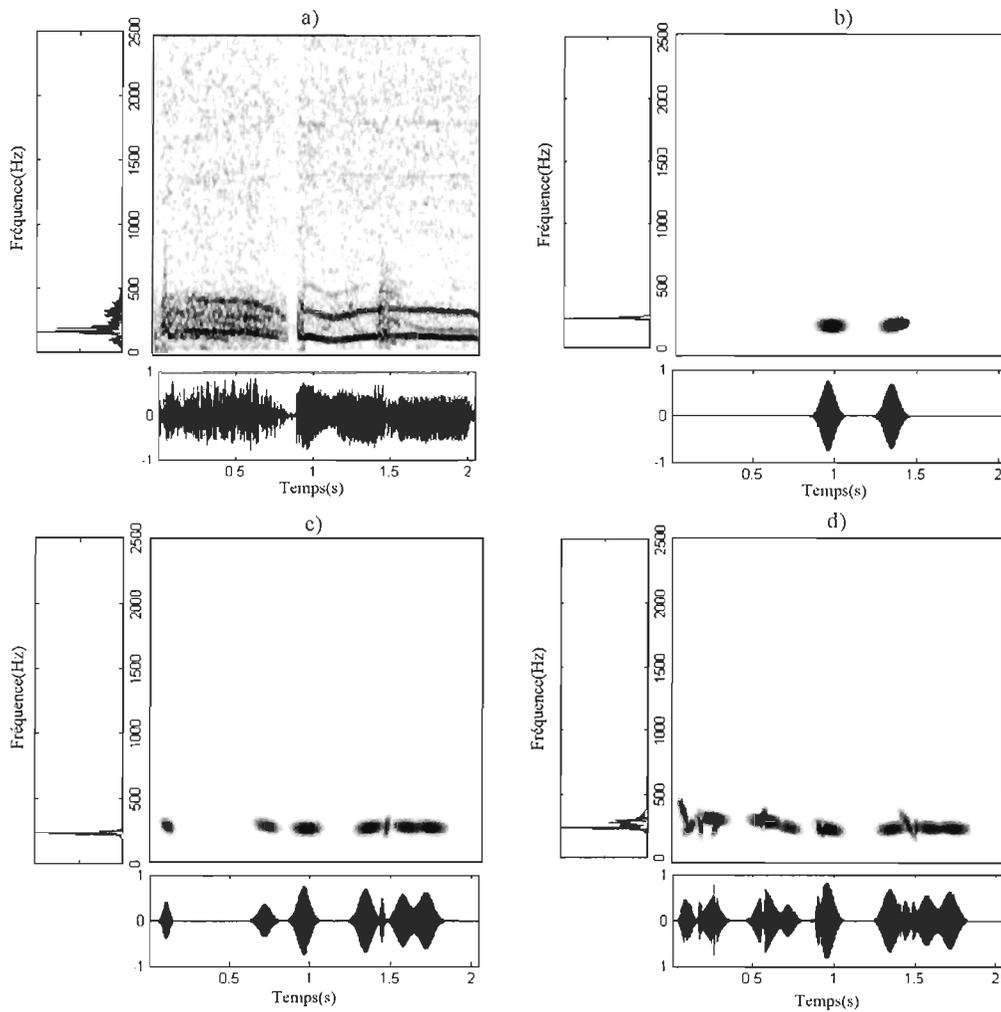


Figure 4.9: Représentation du sibilant par chirplets : (a) le spectrogramme du signal original. (b) 2 chirplets. (c) 7 chirplets. (d) 15 chirplets en utilisant l'outil DiscreteTFDs [76].

4.6.2 Représentation des sibilants par "Chirplets"

Dans cette section, nous appliquons la transformée par *chirplets* aux sibilants. L'idée de base consiste à tenter de modéliser un signal par un nombre limité de chirplets. Pour le même signal utilisé avec les méthodes précédentes, la figure 4.9 montre le résultat de la représentation avec différents nombres de *chirplets*. On constate que *7-chirplets* donnent un meilleur compromis.

La représentation par *chirplets* cherche les composants principaux du signal. Selon le

résultat obtenu, elle détecte seulement les premières harmoniques du sibilant qui ont une énergie plus élevée. Il serait intéressant de l'associer à analyse multibande pour détecter d'éventuels sibilants polyphoniques.

En résumé, cette méthode peut être appliquée à toutes les classes de signaux (continus ou discontinus). Elle extrait les paramètres représentatifs à partir d'un signal multistruktures. Nous pouvons l'employer dans un système de modélisation et de classification. Mais, l'estimation des paramètres est relativement difficile.

CHAPITRE V

RÉSULTATS EXPÉRIMENTAUX

Dans ce chapitre, les méthodes d'analyse des crépitants et des sibilants, décrites précédemment sont testées et évaluées. Deux séries de tests seront réalisées : la première concerne l'évaluation des méthodes de séparation, de détection et de classification des crépitants ; et la seconde concerne la comparaison des techniques de détection ou d'extraction des sibilants.

5.1 Préparation de la base de données

Les données utilisées dans cette recherche proviennent de plusieurs sources : des bases de données ASTRA ^[19] et RALE ^[30], et des fichiers audio accompagnant des livres spécialisés ^[31,32]. Les sons sont catalogués en quatre classes : sons vésiculaires (SV), sibilants (SI), crépitants fins (CF) et gros crépitants (GC). Chaque fichier audio contient un seul cycle respiratoire. Le tableau 5.1 donne la liste des fichiers utilisés. Le nombre de crépitants et la durée de sibilants sont obtenus manuellement et sont considérés comme les valeurs de référence.

Nous ignorons toutes informations sur les sujets (âge, taille, sexe, etc) et les conditions d'enregistrement (le type de capteur, le filtre de antireplissement, le niveau du bruit, etc...). Ce manque d'informations ne constitue pas un handicap, car le but recherché est de réaliser un système de reconnaissance fiable et robuste, qui ne doit surtout pas tenir compte de ses facteurs.

Dans la première série de tests, nous utilisons 18 fichiers : CF5K01-CF5K05, GC5K01-

Tableau 5.1: Base de données utilisée.

Première série de tests (Reconnaissance des crépitants)					Deuxième série de tests (Reconnaissance des sibilants)		
Sons respiratoires	Durées (ms)	Nombre réel des crépitants	Sons respiratoires	Durées (ms)	Sons respiratoires	Durées (ms)	Durées du sibilant (ms)
CF5K01	2164	19	SV5K01	3482	SI8K01	4615	3300
CF5K02	2048	21	SV5K02	3072	SI8K02	4469	3343
CF5K03	2253	18	SV5K03	4452	SI8K03	933	604
CF5K04	2662	14	SV5K04	4091	SI8K04	955	610
CF5K05	9011	43	SI5K01	5939	SI8K05	2096	1338
GC5K01	4710	23	SI5K02	2867	SI8K06	2324	1940
GC5K02	4096	20	SI5K03	1843	SI8K07	1715	941
GC5K03	4096	9	SI5K04	2048	SI8K08	1777	771
GC5K04	3482	8			SI8K09	3454	4160
GC5K05	1434	7			SI8K10	3276	3870
					SI8K11	2047	3484
					SI8K12	2267	3530
Total	35956	182	Total	27794	Total	29928	27891

GC5K05, SV5K01-SV5K04, SI5K01-SI5K04. Les sons respiratoires sont normalisées et échantillonnées à la même fréquence de 5000 Hz. À cause de leur faible puissance, les crépitants sont enregistrés au niveau de la poitrine. Par conséquent, les crépitants sont souvent accompagnés de sons vésiculaires.

Dans la deuxième série, nous utilisons 12 fichiers de sons respiratoires avec sibilants : SI8K01-SI8K12. Ils sont échantillonnés à 8000 Hz et enregistrés au niveau de la trachée. Comparés aux sons vésiculaires, les sons trachéaux sont plus intenses et riches en fréquences, permettant ainsi d'analyser tous les harmoniques des sibilants.

5.2 Critères d'évaluation

Avant d'entamer la comparaison de ces méthodes, nous commençons par définir un certain nombre de critères d'évaluation quantitative.

5.2.1 Taux de séparation

Pour évaluer les performances des méthodes de séparation des crépitants, nous définissons le taux de séparation (TS) par :

$$TS = \frac{NS}{NR} \quad (5.1)$$

où NS représente le nombre de crépitants séparés correctement par le filtre, et NR représente le nombre réel des crépitants.

5.2.2 Rapport signal-sur-bruit (SNR)

Pour évaluer les performances des méthodes de séparation, en terme de préservation de la qualité des sons stationnaires (sons normaux ou avec sibilants), nous définissons le rapport signal-sur-bruit (SNR) par :

$$SNR = 10 \log_{10} \left(\frac{P_X}{P_N} \right) \quad (5.2)$$

où P_X est la puissance du signal d'entrée (normal ou sibilant), et P_N est la puissance du signal récupéré à la sortie non-stationnaire (bruit).

5.2.3 Sensitivité et valeur prédictive positive

La *sensitivité* et la *valeur prédictive positive* sont deux critères souvent utilisés pour évaluer les performances des systèmes de détection et de classification. Elles sont respectivement définies par :

$$\text{Sensitivité} = \frac{\text{Vrais positifs}}{\text{Vrais positifs} + \text{Faux négatifs}} \quad (5.3)$$

$$\text{Valeur prédictive positive} = \frac{\text{Vrais positifs}}{\text{Vrais positifs} + \text{Faux positifs}} \quad (5.4)$$

Dans le cas des crépitants, les "vrais positifs" sont des crépitants qui ont été classifiés crépitants, les "faux négatifs" sont des crépitants qui ont été classifiés comme non-crépitanants, les "vrais négatifs" sont des non-crépitanants qui ont été détectés non-crépitanants, et les "faux positifs" sont des non-crépitanants qui ont été détectés comme crépitants.

5.2.4 Performance de la classification

Pour évaluer les performances des méthodes de classification, nous définissons la performance par le rapport du nombre de crépitants correctement classifiés (NCCC) et du nombre total de crépitants testés (NTCT) :

$$\text{Performance} = \frac{\text{nombre de crépitants correctement classifiés}}{\text{nombre total de crépitants testés}} = \frac{NCCC}{NTCT} \quad (5.5)$$

5.3 Résultats de la séparation des crépitants

Nous avons décrit, au chapitre II, quatre méthodes de séparation des crépitants des sons vésiculaires, basées sur : le filtrage adaptatif (filtre ST-NST), l'inférence floue (filtre

GFST-NST), la transformée en ondelettes (le filtre WTST-NST) et les paquets d'ondelettes (le filtre WPST-NST). Pour évaluer leurs performances, nous utilisons les trois critères suivants : taux de séparation, temps requis et SNR.

Évaluation préliminaire

Dans un premier temps, nous effectuons une comparaison préliminaire des performances de ces techniques. La figure 5.1 donne un exemple de séparation des crépitants par les quatre filtres. Dans cet exemple, ces filtres sont capables de séparer les 3 crépitants fins. En revanche, le filtre ST-NST introduit plus de bruits au début de la séparation ; Le filtrage par GFST-NST n'est pas assez précis, des composantes non-stationnaires sont présentes à la sortie stationnaire et les crépitants sont un peu distordus. Les filtres WTST-NST et WPST-SNT présentent des résultats presque identiques.

En résumé, le filtre ST-NST est une méthode non linéaire de séparation, le seuil ε de la figure 2.3 est défini par une valeur fixe γ et par la densité de probabilité $p(x)$ du signal d'erreur $e(x)$. Dans ce cas, nous avons besoin de définir un ensemble de coefficients (k , γ et le taux d'apprentissage de filtre prédictif) selon le type du signal (crépitant fin, gros crépitant). Le filtre GFST-NST a été entraîné grâce à une banque de données qui contient les résultats du filtre WTST-NST. Il permet d'obtenir presque les mêmes résultats que le filtre WTST-NST mais beaucoup plus rapidement. Pour l'entraînement du système GFST-NST, la base de données est classifiée selon le type du signal (crépitant fin et gros crépitant). Cela signifie que nous avons besoin des informations *a priori* du signal pour choisir le filtre approprié.

Pour ce qui est de l'effet secondaire de la séparation et la robustesse de l'algorithme, les filtres WTST-NST et WPST-NST présentent, évidemment, une meilleure performance. Nous nous limitons, par la suite, à l'évaluation de ces deux filtres. Les sons utilisés sont catalogués en quatre classes : sons vésiculaires (SV5K01, SV5K02, SV5K03, SV5K04), sibilants (SI5K01, SI5K02, SI5K03, SI5K04), crépitants fins (CF5K01, CF5K02, CF5K03, CF5K04, CF5K05) et gros crépitants (GC5K01, GC5K02, GC5K03, GC5K04, GC5K05).

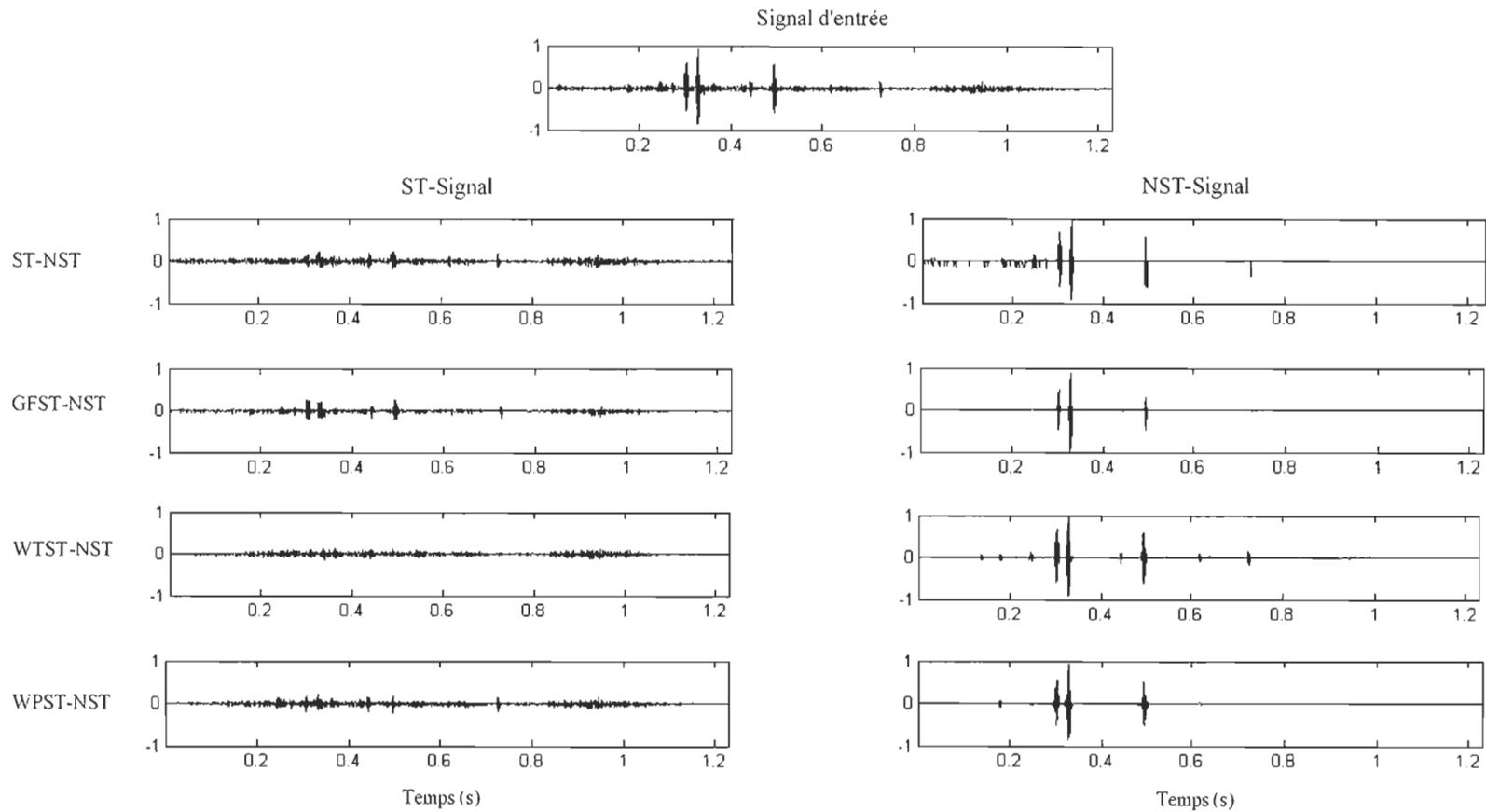


Figure 5.1: Comparaison des performances des quatre filtres de séparation des crépitants. Le signal du haut représente l'entrée des filtres. Les quatre signaux de gauche correspondent aux sorties stationnaires (ST-Signal) des filtres, alors que les quatre de droite correspondent à leurs sorties non-stationnaires (NST-Signal).

Tableau 5.2: Paramètres du filtre WPST-NST et filtre WTST-NST.

	Type d'ondelettes	Longueur de segment	Taux de chevauchement	Profondeur de décomposition	Paramètre 1	Paramètre 2
WPST-NST	Daubechies-8	1024	75%	5	P1=0.75	P2=2
WTST-NST	Daubechies-8	1024	75%	9	$F_{adj} = 3.0$	$\varepsilon = 0.00001$

Les signaux sont divisés en segments de 1024 échantillons ayant un taux de chevauchement de 75%. La transformée par paquets d'ondelettes de notre filtre WPST-NST ^[40] utilise l'ondelette de *Daubechies-8* avec une profondeur de l'arbre de décomposition $j = 5$. Quant au filtre WTST-NST ^[38], il utilise l'ondelette de *Daubechies-8* avec une profondeur de l'arbre de décomposition $j = 9$. Voir le tableau 5.2 pour plus de détails.

Taux de séparation

Afin de mesurer le taux de séparation des crépitants, nous présentons aux filtres uniquement les 10 sons contenant les crépitants (CF5K01–CF5K05 et GC5K01–GC5K05). Les crépitants individuels sont identifiés et localisés manuellement. Nous supposons que les emplacements et le nombre de crépitants obtenus constituent les données de référence pour les filtres de séparation automatique.

La figure 5.2 montre deux exemples de séparation de sons respiratoires en composantes stationnaires (ST-Signal) et non-stationnaires (NST-Signal). Nous indiquons les positions des crépitants par les marques (▼). Dans ce cas, l'ensemble des 14 crépitants fins et celui des 5 gros crépitants sont bien séparés par les deux filtres.

Le tableau 5.3 donne le taux de séparation (TS) et le temps requis (TR) des filtres WPST-NST et WTST-NST pour 10 sons vésiculaires comportant des crépitants. Le taux de séparation du filtre WPST-NST ($TS_{WP} = 98.3 \%$) est comparable à celui du filtre WTST-NST ($TS_{WT} = 98.9 \%$).

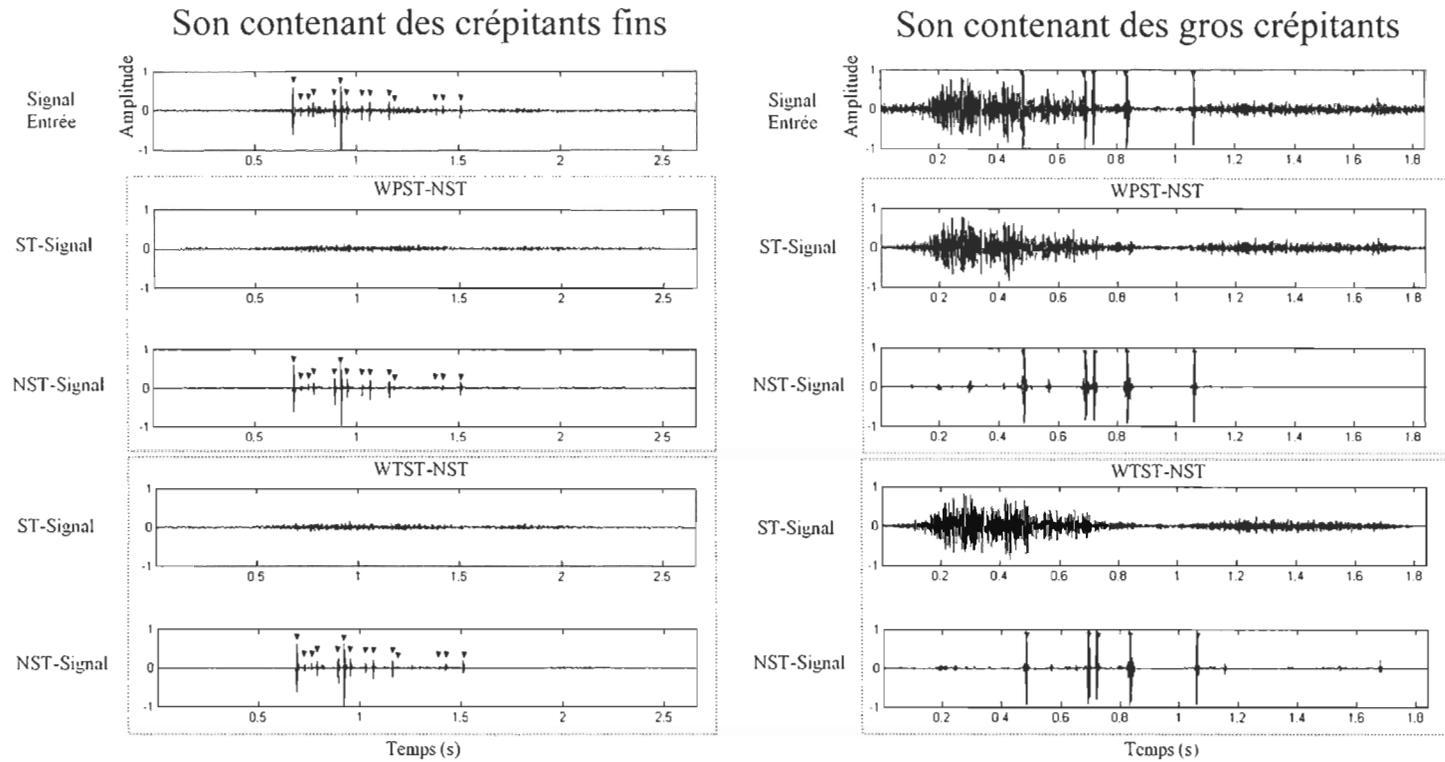


Figure 5.2: Comparaison des performances des filtres WPST–NST et WTST–NST avec des crépitants fins (à gauche) et des gros crépitants (à droite). De chaque côté, le signal du haut représente l’entrée des filtres, les deux signaux du milieu représentent les sorties du filtres WPST–NST, et ceux du bas les sortie du filtre WTST–NST. Les crépitants sont identifiés par les marques (▼).

Tableau 5.3: Taux de séparation (TS) et temps de calcul (TR) des filtres de séparation WPST–NST et WTST–NST, obtenus avec des sons contenant des crépitants fins (CF) et des gros crépitants (GC).

Sons respiratoires	Nombre réel des crépitants	WPST-NST			WTST-NST		
		Nombre de crépitants séparés correctement	Taux de séparation $TS_{WP}(\%)$	Temps requis $TR_{WP}(s)$	Nombre de crépitants séparés correctement	Taux de séparation $TS_{WT}(\%)$	Temps requis $TR_{WT}(s)$
CF01	19	19	100	2.6	19	100	12.5
CF02	21	21	100	2.5	21	100	12.2
CF03	18	17	94.4	2.3	18	100	13.6
CF04	14	14	100	3.3	14	100	16.4
CF05	43	42	97.7	12.1	43	100	57.9
GC01	23	23	100	6.6	22	95.7	31.3
GC02	20	19	95.0	5.4	19	95.0	26.4
GC03	9	9	100	5.3	9	100	25.8
GC04	8	8	100	4.3	8	100	21.7
GC05	7	7	100	1.7	7	100	8.2
Total	182	179	98.3	46.1	180	98.9	226

Temps de calcul

Malgré que les taux de séparation sont comparables (tableau 5.3), notre méthode WPST–NST présente un temps de calcul, ($TR_{WP} = 46.1 s$), 5 fois plus rapide que celui de la méthode WTST–NST ($TR_{WT} = 226 s$). Cette différence est due au fait que l'algorithme WTST–NST est un algorithme itératif, ce qui n'est pas le cas pour WPST–NST.

Quantification du bruit

Pour évaluer le niveau du bruit introduit par les filtres, nous leur présentons uniquement les 8 sons stationnaires qui ne contiennent pas de crépitants (SV5K01–SV5K04 et SI5K01–SI5K04). La figure 5.3 montre les performances des filtres WPST–NST et WTST–NST par deux exemples de sons stationnaires (son vésiculaire et son avec sibilant). Dans ce cas, les signaux présents à la sortie non-stationnaire sont considérés comme du bruit introduit par le filtre. On trouve que le filtre WPST–NST introduit moins de bruit que le filtre WTST–NST.

Le tableau 5.4 donne le niveau du bruit SNR introduit par les deux filtres pour les 4 sons vésiculaires (SV5K01–SV5K04) et 4 sons avec sibilants (SI5K01–SI5K04). Notre filtre WPST–NST présente une meilleure conservation des sons stationnaires que le filtre

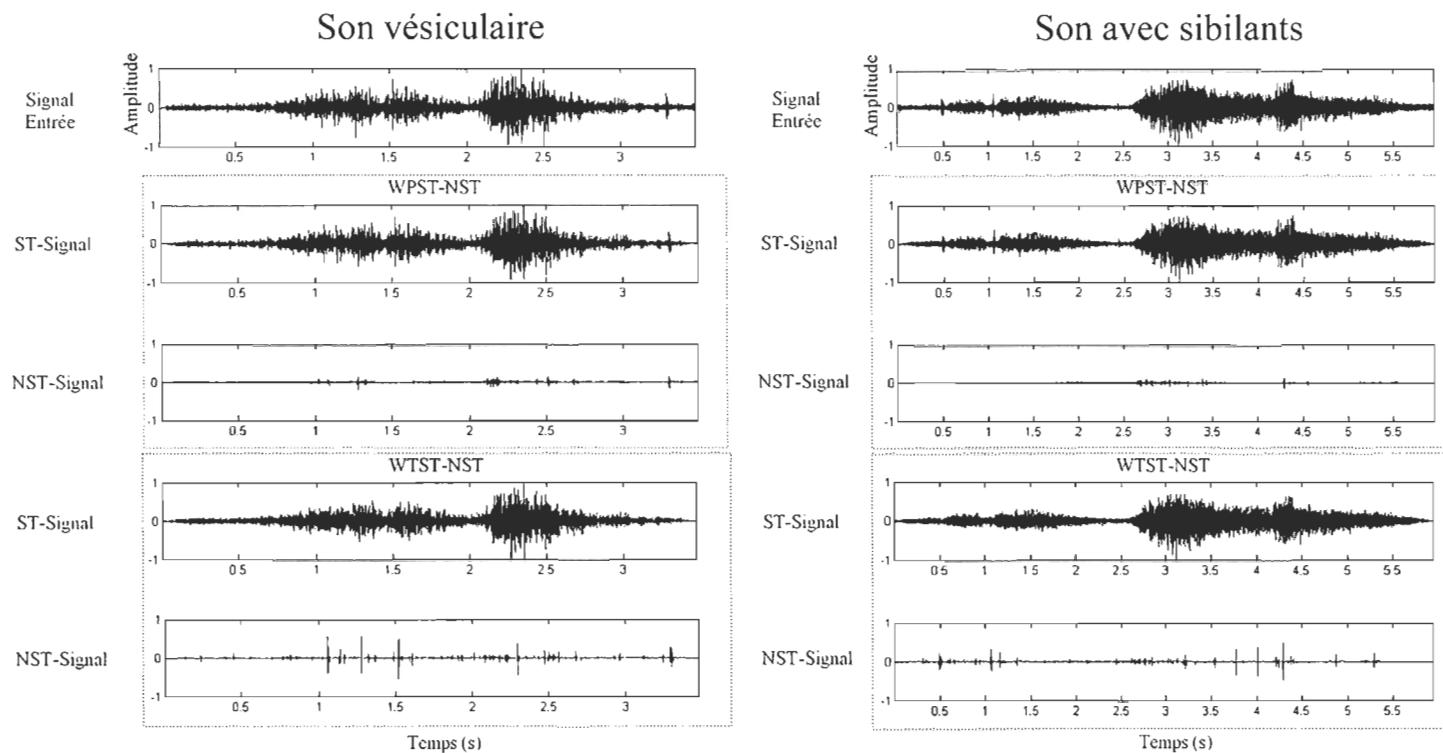


Figure 5.3: Comparaison des performances des filtres WPST–NST et WTST–NST avec un son vésiculaire (à gauche) et un son avec sibilants (à droite) ne contenant aucun crépitant. De chaque côté, le signal du haut représente l’entrée des filtres, les deux signaux du milieu représentent les sorties du filtres WPST–NST, et ceux du bas les sortie du filtre WTST–NST.

Tableau 5.4: Niveau du bruit SNR introduit par les filtres de séparation WPST–NST et WTST–NST. Les tests ont été réalisés avec des sons vésiculaires (SV) et des son sibilants (SI) ne contenant aucun crépitan

	SNR _{WP}	SNR _{WT}		SNR _{WP}	SNR _{WT}
SV01	21.3	14.2	SI01	27.2	17.0
SV02	20.4	13.6	SI02	20.6	21.5
SV03	23.3	18.6	SI03	15.7	19.4
SV04	24.4	20.5	SI04	25.2	14.1
	Sons vésiculaires			Sons avec sibilants	

WTST–NST. En effet, moins de bruit à la sortie non-stationnaire implique que les sons stationnaires sont récupérés entièrement (sans déformation) à la sortie stationnaire du filtre. Il est à noter que, lors de la séparation, il est toujours difficile de trouver un compromis entre la précision de détection et le niveau de bruit introduit à la sortie non-stationnaire.

Ces tests montrent que le filtre WPST–NST présente presque le même taux de séparation (TS) que le filtre WTST–NST, avec un temps de calcul (TR) 5 fois plus faible. Il présente aussi l'avantage d'introduire moins de bruit (SNR élevé). Cela signifie que le filtre WPST–NST fournit une séparation plus propre comparé au filtre WTST–NST.

5.4 Résultats de la détection des crépitants

Nous rappelons que la sortie stationnaire (ST-Signal) et la sortie non-stationnaire (NST-signal) du filtre de séparation sont utilisées par l'étage d'identification des crépitants (voir la figure 1.6). En fait, le signal ST-Signal n'est utilisé que dans le débruitage de NST-signal, où il sert comme référence. La détection d'un crépitan consiste à le localiser dans le temps, c'est-à-dire, connaître l'instant du début et de la fin de ce crépitan. Dans cette section, nous nous limitons à la méthode basée sur la dimension fractale (FD), car elle est de loin la meilleure. Ainsi, nous appliquons ce détecteur aux signaux obtenus par les filtres WPST–NST et WTST–NST, afin de les comparer d'avantage.

Pour évaluer cette technique de détection des crépitants, nous utilisons 18 fichiers sonores, dont 10 sont de type non-stationnaire avec crépitants (CF5K01–CF5K05 et

Tableau 5.5: Résultats de la détection par dimension fractale (algorithme de Sevcik). Les tests ont été réalisés avec les signaux issus des filtres de séparation WPST–NST et WTST–NST

Sons respiratoires	Nombre réel des crépitants	WPST-NST			WTST-NST		
		Vrais positifs	Faux positifs	Faux négatifs	Vrais positifs	Faux positifs	Faux négatifs
CF01	19	18	1	1	18	2	1
CF02	21	20	0	1	21	2	0
CF03	18	17	0	1	17	6	1
CF04	14	14	0	0	14	2	0
CF05	43	39	0	4	42	9	1
GC01	23	22	1	1	22	6	1
GC02	20	17	2	3	18	5	2
GC03	9	8	0	1	9	5	0
GC04	8	8	0	0	8	2	0
GC05	7	6	1	1	5	3	2
SV01	0	0	1	0	0	10	0
SV02	0	0	0	0	0	6	0
SV03	0	0	1	0	0	12	0
SV04	0	0	2	0	0	10	0
SI01	0	0	0	0	0	10	0
SI02	0	0	1	0	0	4	0
SI03	0	0	0	0	0	5	0
SI04	0	0	0	0	0	2	0
Total	182	169	10	13	174	101	8
Sensitivité		$\frac{169}{169+13} = 92.9 \%$			$\frac{174}{174+8} = 95.6 \%$		
Valeur prédictive positive		$\frac{169}{169+10} = 94.4 \%$			$\frac{174}{174+101} = 63.3 \%$		

GC5K01–GC5K05) et 8 de type stationnaire sans crépitants (SV5K01–SV5K04 et SI5K01–SI5K04). Dans cette section, la *sensitivité* et la *valeur prédictive positive* sont utilisées comme critères d'évaluation.

Le tableau 5.5 montre les résultats de l'expérimentation. Le nombre de crépitants réel est semblable au test de séparation, qui est calculé visuellement. Les "faux positifs" indiquent le nombre des fausses détections et les "faux négatifs" indiquent les crépitants qui n'ont pas été détectés (perdus). Ces résultats montrent qu'on obtient pratiquement la même *sensitivité* avec les deux filtres, mais la *valeur prédictive positive* du filtre WPST–NST est beaucoup élevée que celle du filtre WTST–NST. C'est-à-dire que l'on obtient un résultat d'analyse plus fiable avec le filtre WPST–NST. À notre avis, cette performance est due au faible niveau de distorsion introduite par le filtre WPST–NST.

5.5 Résultats de la classification des crépitants

Dans cette section, nous comparons les trois méthodes de classification des crépitants décrites dans le chapitre III. Elles sont basées sur la caractérisation temps-fréquence, le modèle de Prony et la transformée en ondelettes continue. Les crépitants testés proviennent de l'étape de la détection. Nous commençons par enlever les fausses détections et les quelques crépitants fortement distordus. Ensuite, les crépitants individuels sont classifiés manuellement selon l'information *a priori* et les définitions des crépitants fins et des gros crépitants. Le résultat de l'analyse manuelle est considéré comme la "classification de référence". Finalement, chaque crépitant est mis dans un vecteur et marqué en gros ou fin. On dispose de 111 crépitants fins et 54 gros crépitants dans cette base de données.

Pour le test de la classification par transformée en ondelettes continue, le seuil discriminant correspond à l'échelle de dilatation ($s = 3.1$). L'algorithme de classification consiste à trouver l'échelle optimale offrant l'énergie maximale. Pour chaque crépitant, l'énergie est calculée pour différentes valeurs de l'échelle prises dans l'intervalle [1 9] avec un pas de 0.1 (continuité de l'échelle). Les crépitants sont classés en gros ou fin selon que l'échelle optimale est supérieure ou inférieure au seuil ($s = 3.1$). Le résultat de la classification est comparé à la classification manuelle de référence.

Les tests de la détection par caractérisation temps-fréquence et par modèle de Prony sont un peu plus complexes. Dans ces cas, le nombre des paramètres est supérieur à 1, c'est-à-dire que le problème de classification consiste à trouver un discriminant multidimensionnel. Dans ce projet, nous avons opté pour la modélisation par mélange de gaussiennes (GMM) et la méthode "leave one out" pour réaliser la classification.

Nous avons 4 paramètres pour la caractérisation temps-fréquence : GBW, PF, TDW et MDW, et 8 paramètres pour le modèle de Prony : $\{A_1, a_1, f_1, \theta_1\}$ et $\{A_2, a_2, f_2, \theta_2\}$ représentant respectivement l'amplitude, le coefficient d'atténuation, la fréquence et la phase des deux modèles correspondant aux deux moitiés du crépitant. Par exemple, pour tester la méthode de détection à base du modèle de Prony, on construit deux en-

sembles (matrices) à partir de ces paramètres $\{A_1, a_1, f_1, \theta_1, A_2, a_2, f_2, \theta_2\}$ correspondant aux crépitants des deux classes (crépitants fins et gros crépitants). Pour éviter que les données servant à l'apprentissage ne soient pas utilisées lors du test, on divise chacun des deux ensembles en 10 sous-ensembles. Le principe de la méthode "leave one out", consiste à utiliser 9 sous-ensembles de chacune des deux classes pour entraîner les modèles GMM et à prendre le sous-ensemble restant dans chacune des classes pour le tester. Par rotation, l'opération est répétée pour tester les 10 sous-ensembles de chaque classe.

Le tableau 5.6 donne les résultats du test. La classification par transformée en ondelettes continue présente le meilleur résultat, 93.9 % des crépitants sont bien classifiés. Le résultat de la classification par caractérisation temps-fréquence est aussi acceptable, la performance est de 91.5% quand nous prenons les trois paramètres (GBW, PF et MDW). On trouve que la classification ne bénéficie pas du paramètre TDW ; deux raisons peuvent expliquer ce phénomène : 1) la différence de TDW n'est pas évidente entre gros crépitant et crépitant fin ; 2) la façon dont on calcule le TDW n'est pas assez précise. La performance de classification par le modèle Prony est plus faible (72.1 %). Cela est dû au fait que les crépitants sont souvent contaminés par les bruits respiratoires, on ne peut pas réussir à simuler les crépitants distordus avec le modèle Prony.

5.6 Résultat de l'extraction des sibilants

Nous avons décrit, au chapitre IV, quatre méthodes d'extraction des sibilants basées sur le Laplacien d'une gaussienne (LoG), la modélisation auditive (FDDT), la détection d'épisodes de sibilants (WED) et la représentation par "Chirplets". Dans cette section, nous comparerons seulement les deux méthodes (FDDT et WED) qui présentent les meilleures performances. Nous utilisons la deuxième série de fichiers composée de 12 sons respiratoires avec sibilants échantillonnés à une fréquence de 8000 Hz (voir tableau 5.1). Pour calculer le spectrogramme du signal, nous avons utilisé les paramètres de la STFT suivants : fenêtre de type Hanning avec des segments de $L=M=512$ échantillons pour la FDDT, les segments sont de $L=256$ échantillons et la FFT est de longueur $M=2048$ pour

Tableau 5.6: Comparaison des performances des méthodes de classification des crépitants.

Méthode d'extraire les caractéristiques	Paramètres extraits	GMM classification (M=2)			GMM classification (M=4)		
		NTCT	NCCC	Performance(%)	NTCT	NCCC	Performance(%)
caractérisation temps-fréquence	[GBW, PF, TDW, MDW]	165	142	86.1	165	143	86.7
caractérisation temps-fréquence	[GBW, PF, MDW]	165	151	91.5	165	144	87.3
Prony modèle	$[A, a, f, \theta]$	165	110	66.7	165	117	70.9
Prony modèle	$[a, f,]$	165	119	72.1	165	117	70.9
		NTCT		NCCC		Performance(%)	
Ondelettes continues	échelle discriminant=3.1	165		155		93.90	

la WED, le taux d'entrelacement est de 50 %.

Pour évaluer les performances de ces méthodes, nous utilisons la durée (en milliseconde) des segments reconnus comme sibilants. Nous commençons par identifier les sibilants manuellement en se basant sur l'écoute du son et la visualisation de son spectrogramme. La durée du sibilant est calculée manuellement et considérée comme la durée réelle de référence.

La *sensitivité* et la *valeur prédictive positive* sont deux variables qui évoluent dans des sens opposés; l'augmentation de l'un entraîne souvent la diminution de l'autre. Le défi consiste à trouver le bon compromis (équilibre) pour les deux variables. Pour comparer les deux méthodes d'extraction des sibilants, on ajuste leurs seuils afin d'obtenir la même *sensitivité*. Dans le cas de la méthode FDDT, le $Seuil_2$ a été défini pour qu'il puisse varier entre 4 dB à 250 Hz et 10 dB à 4000 Hz; alors que pour la méthode WED, le seuil de continuité temporelle Δt_{WED} est de 32 ms, le seuil de la continuité fréquentielle Δf_{WED} est de 23 Hz et la durée minimale d'un sibilant Δt_{sib} est de 100 ms.

Le tableau 5.4 montre les résultats du test. Avec une même *sensitivité* de 66.6 %, la méthode WED obtient une meilleure *valeur prédictive positive* (95.4 %) contre 77.6 % pour la méthode FDDT.

Tableau 5.7: Performances des méthodes FDDT ^[43] et WED ^[44] d'extraction des sibilants

Sibilants	Total de la durée (ms)	Méthode FDDT					Méthode WED				
		Vrais Positifs (ms)	Faux Positifs (ms)	Faux Négatifs (ms)	Sensitivité (%)	Valeur prédictive positive (%)	Vrais Positifs (ms)	Faux Positifs (ms)	Faux Négatif (ms)	Sensitivité (%)	valeur prédictive positive (%)
SS01	3300	2688	352	612	81.4	88.4	2342	0	958	71.0	100
SS02	3343	2784	896	559	83.3	75.7	2092	258	1251	62.6	89.0
SS03	604	480	96	124	79.5	83.3	584	0	20	96.7	100
SS04	610	256	128	354	42.0	66.7	551	0	59	90.3	100
SS05	1338	580	416	758	43.3	58.2	738	321	600	55.2	69.7
SS06	1940	736	1088	1204	37.9	40.4	1636	128	304	84.3	92.7
SS07	941	640	288	301	68.0	69.0	510	0	431	54.2	100
SS08	771	672	128	99	87.2	84.0	321	0	450	41.6	100
SS09	4160	2688	768	1472	64.6	77.8	1642	128	2518	39.4	92.8
SS10	3870	2976	384	894	76.9	88.6	1451	0	2419	36.6	100
SS11	3484	1664	0	1820	47.8	100	2823	0	661	81.0	100
SS12	3530	2654	32	876	75.2	98.8	3048	0	482	86.3	100
Moyenne					66.6%	77.6%				66.6%	95.4%

CONCLUSION

Notre projet de recherche consiste à construire un système de reconnaissance automatique des sons respiratoires adventices. Nous nous sommes particulièrement intéressés aux crépitants fins et aux gros crépitants, qui sont considérés comme discontinus ou transitoires, et aux sibilants qui sont considérés comme continus ou quasi-stationnaires. Ce projet entre dans le cadre d'une problématique de recherche plus vaste, qui vise à développer des techniques de reconnaissance des sons adventices dans le but de réaliser un système d'aide au diagnostic des maladies pulmonaires.

Notre démarche se résume en trois étapes : 1) la séparation des sons respiratoires en composantes stationnaires/non-stationnaires, 2) l'identification des crépitants et 3) la détection des sibilants. Dans le chapitre II, nous avons décrit trois techniques de séparation basées sur le filtrage adaptatif (ST-NST), le système d'inférences floue (GFST-NST), et la transformée en ondelettes (WTST-NST), ainsi qu'une nouvelle technique basée sur les paquets d'ondelettes (WPST-NST), que nous avons proposée récemment. Dans le chapitre III, nous avons décrit les paramètres de détection des crépitants basés sur l'enveloppe de l'amplitude absolue (EA), l'enveloppe de l'énergie (EE), et la dimension fractale (FD) avant de présenter les méthodes de classification des crépitants basées sur l'analyse temps-fréquence, modèle de Prony et la transformée en ondelettes continue. Dans le chapitre IV, nous avons discuté de quatre méthodes de détection des sibilants basées sur : le Laplacien d'une gaussienne (LoG), la modélisation auditive (FDDT), la détection d'épisodes de sibilants (WED) et la représentation par "Chirplets". Enfin, les performances des méthodes discutées sont fournies au chapitre V.

La base de données utilisée comprend deux séries d'enregistrements. La première série, de 18 enregistrements, a été utilisée pour évaluer les méthodes de séparation et d'identification des crépitants, alors que la deuxième, de 12 enregistrements, a été utilisée pour évaluer les techniques de détection des crépitants. Nous présentons le bilan de chaque étape de l'expérimentation.

– Séparation des crépitants

Le filtre de séparation ST-NST est le plus simple et le rapide, mais il n'est pas assez précis ; le filtre GFST-NST fonctionne mieux dans les conditions limites, mais il a besoin d'informations *a priori* sur le signal ; comparativement, les filtres WTST-NST et WPST-NST présentent les meilleures performances. Pour un taux de séparation relativement identique (98 %), le filtre WPST-NST est 5 fois plus rapide que filtre WTST-NST. En plus, il introduit moins de bruit durant la séparation, donc, moins de déformation. En résumé, le filtre WPST-NST est de loin le plus avantageux.

– Détection des crépitants

La détection des crépitants est réalisée en utilisant le paramètre dimension fractale (FD). Ce détecteur est capable de trouver, avec précision, les positions (les frontières) des crépitants qu'ils soient faibles ou entrelacés. Pour les tests réalisés sur les signaux obtenus aux sorties non-stationnaires des filtres WPST-NST et WTST-NST, les taux de détection positive (*sensitivité*) sont respectivement de 92.9 % et 95.6 %, alors que les valeurs prédictives positives sont respectivement de 94.4 % et 63.3 %. Les meilleures performances obtenues avec notre filtre WPST-NST sont dues probablement à son faible niveau de distorsion.

– Classification des crépitants

Pour classifier les crépitants en crépitants fins et gros crépitants, nous avons expérimenté les trois méthodes citées ci-dessus. La classification par transformée en ondelettes continue fournit la meilleure performance (93.9 %) mais avec un temps de calcul plus important. La classification par trois paramètres temps-fréquence (GBW, PF et MDW) constitue un bon compromis, la mesure des paramètres est simple et le résultat est acceptable (*performance*=91.5 %). Les paramètres du modèle de Prony ne fonctionnent pas correctement dans notre expérimentation (*performance*=72.1 %).

– Détection des sibilants

Pour la détection des sibilants, nous avons testé et comparé les quatre méthodes listée ci-dessus. La méthode basée sur les *chirplets* extrait les principaux harmo-

nique du signal. Dans notre cas, elle trouve seulement le premier harmonique du sibilant, et l'estimation de ses paramètres est lourde en temps de calcul. Les trois autres méthodes basées sur le spectrogramme donnent relativement de bonnes performances. Avec les pré-traitements et les différentes définitions du seuil, la méthode basée sur la détection d'épisodes de sibilants (WED) donne la meilleure performance, suivie par celle basée sur la modélisation auditive (FDDT). Avec une même *sensitivité* de 66.6 %, la méthode WED obtient une meilleure *valeur prédictive positive* (95.4 %) contre 77.6 % pour la méthode FDDT.

En résumé, une grande partie de ce mémoire vise l'analyse des signaux discontinus. Elle présente une méthode robuste, efficace et complètement automatique pour extraire et classifier les crépitants. Pour ce qui est de l'analyse des sibilants, nous avons abordé le problème de la détection par les techniques basées sur le spectrogramme, il reste encore à approfondir cette démarche dans le but de finaliser ce module.

BIBLIOGRAPHIE

- [1] Santé Canada. Les maladies respiratoires au Canada. <http://www.hc-sc.gc.ca/pphb-dgspsp/publicat/rdc-mrc01/pdf/mrc0901f.pdf>, Septembre 2001.
- [2] A.R.A. Sovijärvi, L.P. Malmberg, G. Charbonneau, J. Vanderschoot, F. Dalmasso, C. Sacco, M. Rossi, and J.E. Earis. Characteristics of breath sounds and adventitious respiratory sounds. *European Respiratory Review*, 10 :591–596, 2000.
- [3] E. Lens, G. Postiaux, and P. Chappelle. Application in bedside medicine of automated spectral analysis of breath sound, wheezing and crackles. In *Proceedings of the 12th International Lung Sounds Conference*, CNRS, Paris, 15-19 February 1987.
- [4] M. Nissan and N. Gavriely. A microcomputer based lung sounds analysis. *Comput. Methods Programs Biomed.*, 40(1) :7–13, 1993.
- [5] X. Sun, B.M. Cheetham, and J.E. Earis. Real time analysis of lung sounds. *Technology and Health Care*, 6(1) :3–10, 1998.
- [6] A.R.A. Sovijärvi, P. Helistö, L.P. Malmberg, K. Kallio, E. Paaanen, A. Saarinen, P. Lipponen, S. Haltsonen, L. Pekkanen, P. Piirilä, L. Näveri, and T. Katila. A new versatile PC-based lung sound analyzer with automatic crackle analysis (HeLSA). *Technology and Health Care*, 6(1) :11–22, 1998.
- [7] M. Waris, P. Helistö, S. Haltsonen, A. Saarinen, and A.R. Sovijärvi. A new method for automatic wheeze detection. *Technol Health Care*, 6(1) :33–40, 1998.
- [8] L. Pesu, P. Helistö, E. Adomovic, J.C. Pesquet, A. Saarinen, and A.R.A. Sovijärvi. Classification of respiratory sounds based on wavelet packet decomposition and learning vector quantization. *Technology and Health Care*, 6(1) :65–74, 1998.
- [9] S. Rietveld, M. Oud, and E.H. Dooijes. Classification of asthmatic breath sounds : preliminary results of the classifying capacity of human examiners versus artificial neural networks. *Comput. Biomed. Res.*, 32 :440–448, 1999.

- [10] L.R. Waitman, K.P. Clarkson, J.A. Barwise, and P.H. King. Representation and classification of breath sounds recorded in an intensive care setting using neural networks. *Journal of Clinical Monitoring and Computing*, 16 :95–105, 2000.
- [11] Z. Dokur and T. Olmez. Classification of respiratory sounds by using an artificial neural network. *International Journal of Pattern Recognition and Artificial Intelligence*, 17(4) :567–580, 2003.
- [12] A. Kandaswamy, C. Sathish Kumar, Rm. Pl. Ramanathan, S. Jayaraman, and N. Malmurugan. Neural classification of lung sounds using wavelet coefficients. *Computers in Biology and Medicine*, 34 :523–537, 2004.
- [13] E.C. Guler, B. Sankur, Y.P. Kahya, and S. Raudys. Two-stage classification of respiratory sound patterns. *Computers in Biology and Medicine*, 35(1) :67–83, 2005.
- [14] M. Bahoura and C. Pelletier. New parameters for respiratory sound classification. In *Electrical and Computer Engineering, 2003. IEEE CCECE 2003. Canadian Conference on*, volume 3, pages 1457–1460, Montreal, Canada, May 4-7 2003.
- [15] M. Bahoura and C. Pelletier. Respiratory sounds classification using gaussian mixture models. In *Electrical and Computer Engineering, 2004. IEEE CCECE 2004. Canadian Conference on*, volume 3, pages 1309–1312, Niagara Falls, Canada, May 2-5 2004.
- [16] M. Bahoura and C. Pelletier. Respiratory sounds classification using cepstral analysis and gaussian mixture models. In *26th Annual Conference of the IEEE EMBS*, pages 9–12, San Francisco, California, September 1-5 2004.
- [17] C. Pelletier. Classification des sons respiratoires en vue d’une détection automatique des sibilants. Master’s thesis, Université du Québec à Rimouski, Canada, 2006.
- [18] P. Marquet. Les bruits d’origine ventilatoire. *Rev. Mal. Resp.*, 12 :87–101, 1995.
- [19] L’auscultation pulmonaire. CD audio des laboratoires Astra France. <http://www.ifrance.com/y-detauriac/cdrom/pneumo.htm>.
- [20] P. Piirila and A.R.A. Sovijarvi. Crackles : recording, analysis and clinical significance. *European Respiratory Review*, 8 :2139–2148, 1995.

- [21] B. Karnath and M.C. Boyars. Pulmonary auscultation. *Hopital Physician*, pages 22–26, 2002.
- [22] G. Charbonneau, E. Ademovic, B.M.G. Cheetham, L.P. Malmberg, J. Vanderschoot, and A.R.A. Sovijarvi. Basic techniques for respiratory sound analysis. *European Respiratory Review*, 10 :625–635, 2000.
- [23] R. Mikami, M. Murao, D.W. Cugell, J. Chretien, P. Cole, J. Meier-Sydow, R.L. Murphy, and R.G. Loudon. International symposium on lung sounds. synopsis of proceedings. *Chest*, 92 :342–345, 1987.
- [24] R.G. Loudon and R.L. Murphy. State of the art : lung sounds. *Am.Rev.Respir.Dis.*, 130 :663–673, 1984.
- [25] Y. Takezawa, F. Shirai, and S. Sawaki et al. Comparaison of wheezes over the trachea and on the chest wall. In *the Fifth International Lung Sounds Conference*, London, England, 1980. International Lung Sound Association.
- [26] T.R Fenton, H. Pasterkamp, A. Tal, and V. Chernick. Automated spectral characterisation of wheezing in asthmatic children. *IEEE Trans. Biomed. Eng.*, BNE-32(1) :50–55, 1985.
- [27] A.R.A. Sovijärvi, F. Dalmasso, J. Vanderschoot, L.P. Malmberg, G. Righini, and S.A.T. Stoneman. Definition of terms for applications of respiratory sounds. *European Respiratory Review*, 10 :597–610, 2000.
- [28] B. Sankur, E.C. Güler, and Y.P. Kahya. Multiresolution biological transient extraction applied to respiratory crackles. *Comp. Biol. Med.*, 26 :25–39, 1996.
- [29] P. Mastorocostas and J. Theocharis. A dynamic fuzzy-neural filter for separation of discontinuous adventitious sounds from vesicular sounds. *Computers in Biology and Medicine*, 57 :60–69, 2007.
- [30] The R.A.L.E lung sound 3.0. <http://www.rale.ca>, 2003.
- [31] S. Lehrer. *Understanding Lung Sounds, 3/e (Booklet with Audio CD)*. W.B. Saunders Company, 2002.

- [32] R. Wilkins, J. Hodgkin, and B. Lopez. *Fundamentals of Lung and Heart Sounds, 3/e (Book and CD-ROM)*. C.V. Mosby, 2004.
- [33] Y.P. Kahya and CA. Yilmze. Modeling of respiratory crackles. In *Proceedings of the 22nd annual EMBS international confernece*, pages 23–28, July 2000.
- [34] X. Lu and M. Bahoura. An automatic system for crackles detection and classification. In *Electrical and Computer Engineering, 2006. IEEE CCECE 2006. Canadian Conference on*, Ottawa, Canada, May 7-10 2006.
- [35] M. Ono, K. Arakawa, M. Mori, T. Sugimoto, and H. Harashima. Separation of fine crackles from vesicular sounds by a nonlinear digital filter. *IEEE Trans Biomed Eng*, 36(2) :286–291, 1989.
- [36] K. Arakawa, H. Harashima, M. Ono, and M. Mori. Non-linear digital filters for extracting crackles from lung sounds. *Front. Med. Biol. Eng.*, 3(4) :245–257, 1991.
- [37] L.J. Hadjileontiadis and S.M. Panas. Nonlinear separation of crackles and squawks from vesicular sounds using third-order statistics. In *18th International Conference of the IEEE Engineering in Medicine and Biology Society*, volume 5, pages 2217–2219, 1996.
- [38] L.J. Hadjileontiadis and S.M. Panas. Separation of discontinuous adventitious sounds from vesicular sounds using a wavelet-based filter. *IEEE Trans Biomed Eng*, 44(12) :1269–1281, 1997.
- [39] Y.A.Tolias, L.J. Hadjileontiadis, and S.M. Panas. Real-time separation of discontinuous adventitious sounds from vesicular sounds using a fuzzy rule-based filter. *IEEE Trans Inf Technol Biomed*, 2(3) :204–215, 1998.
- [40] M. Bahoura and X. Lu. Separation of crackles from vesicular sounds using wavelet packet transform. In *Proceedings of the International Conference on Acoustics, Speech and Signal Processing (ICASSP 2006)*, Toulouse, France, May 2006.
- [41] L.J. Hadjileontiadis and I.T. Rekanos. Detection of explosive lung and bowel sounds by means of fractal dimension. *Signal Processing Letters, IEEE*, 10(10) :311–314, 2003.

- [42] R. J. Riella, P. Nohama, R. F. Borges, and A. L. Stelle. Automatic wheezing recognition in recorded lung sounds. In *Proceedings of the 25th Annual International Conference of the IEEE EMBS*, pages 17–21, September 2003.
- [43] Y. Qiu, A. R. Whittaker, M. Lucas, and K. Anderson. Automatic wheeze detection based on auditory modelling. In *Proc. Inst. Mech. Eng.*, pages 219–27, May 2005.
- [44] S. A. Taplidou, L. J. Hadjileontiadis, T. Penzel, V. Gross, and S. M. Panas. WED : An efficient wheezing-episode detector based on breath sounds spectrogram analysis. In *Proceedings of the 25th Annual International Conference of the IEEE EMBS*, pages 17–21, September 2003.
- [45] J.C. O’Neill, P. Flandrin, and W.C. Karl. Sparse representations with chirplets via maximum likelihood estimation. <http://citeseer.ist.psu.edu/389174.html>.
- [46] R. Chassaing. *DSP : Applications Using C and the TMS320C6x DSK*. Wiley Interscience, 2002.
- [47] F. Chevré and F. Guély. La logique floue. Technical Report CT-191, Groupe Schneider, http://www.schneider-electric.com/cahier_technique/fr/pdf/ct191.pdf, mars 1998.
- [48] J.S.R. Jang. ANFIS : Adaptive-Network-based Fuzzy Inference Systems. *IEEE Transactions on Systems, Man, and Cybernetics*, 23 :665–684, May 1993.
- [49] M. Sugeno. *Industrial applications of fuzzy control*. Elsevier Science Pub. Co., 1985.
- [50] M. Bahoura. *Analyse des Signaux Acoustiques Respiratoires : Contribution à la Détection Automatique des Sibilants par Paquets d’Ondelettes*. PhD thesis, Thèse de Université de Rouen, Janvier 1999.
- [51] M. Bahoura. Réduction du bruit par ondelettes : Application au rehaussement de la parole. Technical report, ERMETIS, Université du Québec à Chicoutimi, Mai 2000.
- [52] J. Landré and F. Truchetlet. Approche exploratoire multirésolution basée sur le contenu d’une base d’images paléontologique. pages 2013–2020. Proceedings of Information Processing and Management of Uncertainty (IPMU), Annecy, Annecy, July 2002.

- [53] S. Mallat. A Theory for Multiresolution Signal Decomposition : The Wavelet Representation. *IEEE Trans. Patt. Anal. Mach. Intell*, 11 :674–693, July 1989.
- [54] R.R. Coifman and M.V. Wickerhauser. Best-adapted wave packet bases. Preprint, Yale University, 1990.
- [55] M. V. Wickerhauser. INRIA lectures on wavelet packet algorithms. In Pierre-Louis Lions, editor, *Problèmes Non-Linéaires Appliqués, Ondelettes et Paquets D'Ondes*, pages 31–99. INRIA, Roquencourt, France, 17–21 June 1991. Minicourse lecture notes.
- [56] R.R. Coifman and M.V. Wickerhauser. Entropy-based algorithm for best basis selection. *IEEE Trans. Inform. Theory*, 38 :713–718, 1992.
- [57] L. Vannuccini, M. Rossi, and G. Pasquali. A new method to detect crackles in respiratory sounds. *Technol Health Care*, 6(1) :75–9, 1998.
- [58] T. Kaisla, A. Sovijarvi, P. Piirila, H. M. Rajala, S. Haltsonen, and T. Rosqvist. Validated method for automatic detection of lung sound crackles. *Med Biol Eng Comput*, 29(5) :517–21, 1991.
- [59] Wikipedia foundation. <http://en.wikipedia.org/wiki/Dimension>.
- [60] K. J. Falconer. *Fractal Geometry : Mathematical Foundations and Applications*. John Willey and Sons, 2003.
- [61] B. B. Mandelbrot. *The Fractal Geometry of Nature*. W. H. Freeman and Company, 1982.
- [62] M. J. Katz. Fractals and the analysis of waveforms. *Computers in Biology and Medicine*, 18(3) :145–156, 1988.
- [63] C. Sevcik. A procedure to estimate the fractal dimension of waveforms. *Complexity International [Online]*, Available from : <http://journal-ci.csse.monash.edu.au/ci/vol05/sevcik/sevcik.html>, 5, 1998. Visited March 15, 2007.
- [64] J. Hoevers and R. G. Loudon. Measuring crackles. *Chest*, 98 :1240–3, November 1990.

- [65] M. Munakata, H. Ukita, I. Doi, Y. Ohtsuka, Y. Masaki, Y. Homma, and Y. Kawakami. Spectral and waveform characteristics of fine and coarse crackles. *Thorax*, 46(9) :651–657, September 1991.
- [66] M.Y. Buruk, Y.P. Kahya, and B. Sankur. Clustering analysis of respiratory crackles. In *Biomedical Engineering Days, 1998. Proceedings of the 1998 2nd International Conference*, pages 137–138, 20-22 May 1998.
- [67] M. Du, F.H.Y. Chan, F.K. Lam, and J. Sun. Crackle detection and classification based on matched wavelet analysis. In *Engineering in Medicine and Biology society, 1997. Proceedings of the 19th Annual International Conference of the IEEE*, volume 4, pages 1638–1641, 30 Oct-2 Nov, 1997.
- [68] R.P. Baughman and R.G. Loudon. Quantitation of wheezing in acute asthma. *Chest*, 86(5) :718–722, 1984.
- [69] B. R. Glasberg and B. C. J. Moore. Derivation of auditory filter shapes from notched-noise data. In *Hearing Research*, pages 103–138, August 1990.
- [70] C. M. Reed and R. C. Bilger. A comparative study of S/No and E/No. *J. Acoust. Soc. Am.*, 53 :1039–1044, 1973.
- [71] A. Rosenfeld and P. Pfaltz. Sequential operations in digital picture processing. In *Journal of the Association for Computing Machinery*, pages 471–494, 12 1966.
- [72] S. Mann and S. Haykin. The chirplet transform : A generalization of Gabor’s logon transform. *Vision Interface '91*, pages 205–212, June 3-7 1991.
- [73] S. Mann and S. Haykin. The Adaptive Chirplet : An Adaptive Wavelet Like Transform. *SPIE, 36th Annual International Symposium on Optical and Optoelectronic Applied Science and Engineering*, 21-26 July 1991.
- [74] J. Cui and W. Wong. The adaptive chirplet transform and visual evoked potentials. *Biomedical Engineering, IEEE Transactions on*, 53(7) :1378–1384, July 2006.
- [75] S. Mann and S. Haykin. The chirplet transform : Physical considerations. *IEEE Trans. Signal Processing*, 43(11) :2745–2761, November 1995.

[76] DiscreteTFDs : Time-Frequency Analysis Software. <http://tfd.sourceforge.net/>.

Index

- Analyse multirésolution, 31
- Base de données, 10, 80
- Calcul du spectrogramme, 63
- Chirplets, 76
- Classification des crépitants, 52
- Crépitant, 8, 10, 12
- Détection des crépitants, 45
- Detection des contours, 65
- Dimension fractale, 47
- Durée de la déflexion initiale, 54
- Durée de la déflexion maximale, 54, 55
- Durée de la déflexion totale, 54, 55
- Durée des deux premiers cycles, 8, 54
- Episodes des sibilants, 73
- Extraction des sibilants, 93
- Filtre
 - de débruitage, 44
 - de séparation, 17, 29, 34
 - gaussien, 65
 - non-linéaire, 16
 - prédictif, 16
- Filtres à miroirs quadratiques, 33
- Fonction d'appartenance, 22
- Fréquence du pic, 9, 55
- Inférence floue, 25
- Laplacien d'une gaussienne, 64
- Largeur de la déflexion initiale, 8
- Logique floue, 20
- Modélisation auditive, 68
- Modélisation par mélange de gaussiennes,
 - 92
- Modèle de Prony, 55
- Nomenclature des sons respiratoires, 3
- Opérateurs de logique floue, 24
- Paquets d'ondelettes
 - Recherche de la meilleure base, 38
 - Transformée, 37
- Quantification du bruit, 88
- Résolution
 - fréquentielle, 33, 63
 - temporelle, 33, 63
- Règles d'inférence, 24

Rapport signal-sur-bruit, 82

Ronchus, 7, 10

Sensitivité , 83

Sibilant, 7, 10, 14

- monophonique, 7
- polyphonique, 7

SNR, 82

Sons non-stationnaires, 15

Sons respiratoires

- adventices continus, 7
- adventices discontinus, 8

Sons respiratoires normaux, 4

Sons stationnaires, 15

Spectrogramme, 63

Système d'inférence floue, 25

Taux de séparation, 82, 86

Transformée de Fourier à court terme, 63

Transformée en ondelettes

- Continue, 30, 58
- Discrete, 31

Valeur prédictive positive, 83

Rapport-Gratuit.com