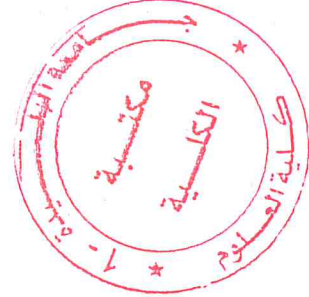


MA-004-441-1

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université Saad Dahlab Blida 1

N° D'ordre :



Faculté des sciences

Département d'informatique

Mémoire Présenté par :

FEDDOUS Asmaa

ELKHEIR Sabrina

En vue d'obtenir le diplôme de master

Domaine : Mathématique et informatique

Filière : Informatique

Spécialité : Informatique

Option : Ingénierie de logiciel

Sujet :

Détection des coups de feu en vue de la mise en œuvre d'un système de surveillance audio

Soutenu le :

Président
Examineur
Examineur

Mr Cherif Zaher .

Promoteur : Dr. Farid YKHLEF, Université SAAD DAHLEB, BLIDA.

Encadreur : Dr. Fayçal YKHLEF, CDTA, Alger.

Promotion
2016 / 2017

MA-004-441-1

Remerciements

Louanges à Dieu le tout Puissant l'Omnipotent qui nous a fait la faveur de nous donner le courage, la force et la volonté pour mener à bien ce modeste travail.

Nos premiers remerciements sont adressés à notre promoteur, Dr. Fayçal YKHLEF, Maître de Recherche B au CDTA, qui nous a proposé ce thème et nous a fait confiance malgré nos connaissances plutôt légères dans le domaine de traitement de signal. Nous le remercions pour son aide précieuse et ses judicieux conseils et pour le temps qu'il nous a consacré.

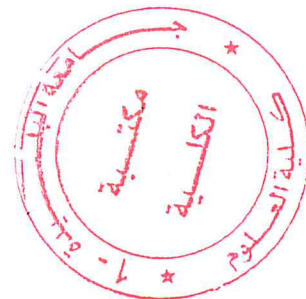
Nous remercions la Direction du Centre de Développement des Technologies Avancées C.D.T.A et en particulier les dirigeants de la Division Architecture des Systèmes et Multimédias de nous avoir accueilli et mis à notre disposition les conditions favorables pour la réalisation de ce présent travail.

Nous remercions aussi notre promoteur, Dr. YKHLEF Farid, Maître de Conférence A, enseignant chercheur au département d'électronique de l'université BLIDA 1, pour sa disponibilité et son aide.

Nous adressons notre très grand respect aux membres du jury qui nous ont fait l'honneur d'évaluer ce modeste de travail.

Nos sincères remerciements s'adressent aussi à tous les enseignants de la faculté des sciences de l'université BLIDA 1, surtout ceux du département "INFORMATIQUE."

Enfin, nous tenons à exprimer nos remerciements envers toute personne qui a contribué de près ou de loin à la réalisation de ce modeste travail.



DEDICACE

Je dédie ce mémoire :

À Ma très chère mère Nacira.

*Tu représentes pour moi la source de tendresse et l'exemple du Dévouement qui n'a pas
Cessé de m'encourager et de prier pour moi. Aucune dédicace ne saurait être assez
Éloquente pour exprimer ce que tu mérites pour tous les sacrifices que tu n'as cessés de me
donner depuis ma naissance.*

À Mon Père Menouer.

*Qui m'a indiqué la bonne volonté, Ce travail est le fruit de tes sacrifices que tu as consentis
Pour mon éducation et ma formation. Aucune dédicace ne saurait exprimer l'amour,
L'estime, et le respect que j'ai toujours eu pour vous.*

*À mes frères **Katib Sifeddine, Chakib** et mes sœurs **Amina Imane et Meriem**.*

*À mes amies : **Wassila, Ikram, Khanssa et Sarah**.*

Et tous les enseignants qui m'ont enseigné chacun son nom

À tous ceux qui j'aime et qui m'aiment

*À mon meilleur binôme **Sabrina** et toute sa famille*

Sans oublier la promotion IL2017

Asmaa

A mes très chères parents Mohamed et Houria

C'est avec profonde gratitude et l'amour que je vous porte et s'incères mots, je mets entre vos main le fruit de longues années d'études, de longs mois de votre tendresse, de longs jours d'apprentissage.

Votre soutien et votre encouragement m'ont toujours donne de la force pour persévérer et pour prospérer dans la vie.

Chaque ligne de ce mémoire chaque mot et chaque lettre vous exprime la reconnaissance, le respect, l'estime et le merci d'être mes parents.

A mes grands mères Yaya (Baya) et djeda (Houria)

Je vous dédie ce travail en témoigne de gratitude d'estime et d'attachement. Puisse Allah vous accorder santé, langue vie et prospérité.

A mes frères et mes sœurs

*Avec tous mes vœux de les voir réussir dans leurs vies.
Amon petit ange^Wassim.*

A mes amies

*A qui je souhaite le succès, en le remerciant pour l'amitié qui nous a toujours unis.
A tous ceux qui me sont chers.*

A tous nos professeurs qui nous ont enseignés.

Et pour finir, a ma chère amie Asma avec qui j'ai partagée des moments spéciaux et à qui je souhaite la réussite et le bonheur ainsi qu'à toute sa famille.

ELKHEIR Sabrina.

ملخص

الهدف من هذا العمل هو تصميم وحدة الكشف عن تبادلات اطلاق النار من أجل تنفيذ نظام المراقبة الصوتية. حيث نركز في هذه المذكرة على تحديد الصوت الذي ينجم عن اطلاق النار. تم دراسة حلين مختلفين. يتكون الحل الأول من مرحلة واحدة و هي المعالجة. وهو يتألف من استخدام تحويل الموجات المستمرة لوضع قواعد القرار. ويتكون الحل الثاني من مرحلتين من المعالجة، استخراج الصفات والتصنيف. ويستند استخراج السمات أيضا على استخدام تحويل الموجات المستمرة لتوليد عدة قياسات الطاقة. وتستند مرحلة التصنيف إلى آلات دعم الناقلات. تم استخراج مجموعة الصوت المستخدمة في تجاربنا من قاعدة بيانات سونودوغ. مقاييس القياس المستخدمة في عملنا هي: المعدل الحقيقي الايجابي (م ح ا)، المعدل الحقيقي السلبي (م ح س)، معدل الكاذب الايجابي (م ك ا)، معدل الكاذب السلبي (م ك س)، و معدل التصنيف الجيد (م ت ج). النتائج التي تم الحصول عليها مرضية.

بالنسبة للنظام الأول، نحصل على النتائج التالية: م ت ج = 0.77، م ح ا = 0.873، م ح س = 0.67، م ك ا = 0.33، م ك س = 0.126. اما بالنسبة للنظام الثاني لدينا م ت ج = 0.983، م ح ا = 1، م ح س = 0.967، م ك ا = 0.033، م ك س = 0.

الكلمات المفتاحية: تحويل الموجات المستمر، آلات دعم الناقلات ، طلاقات نارية.

Résumé

L'objectif de ce travail est de concevoir un module de détection des coups de feu (CF) en vue de la mise en œuvre d'un système de surveillance audio. Nous nous focalisons dans ce mémoire sur la reconnaissance des CF. Deux systèmes différents ont été étudiés. Le premier système est composé d'un seul étage de traitement. Il consiste en l'utilisation de la transformée en ondelettes continue (TOC) en vue de l'élaboration des règles de décision. Le deuxième système est composé de deux étages de traitement, l'extraction d'attributs et la classification. L'extraction d'attributs est basée aussi sur la TOC pour la génération de plusieurs mesures énergétiques. L'étage de classification est basé sur les Support Vector Machines (SVMs). Le corpus sonore utilisé dans nos expérimentations est extrait à partir de la base de données SOUNDDOG. Les métriques d'évaluations utilisées dans notre travail sont : le taux de vrais positifs (TVP), le taux de vrai négatifs (TVN), le taux de faux positifs (TFP), le taux de faux négatifs (TFN) et le Taux de Bonne Classification (TBC). Les résultats obtenus sont satisfaisants.

Pour le premier système les résultats obtenus sont : TBC=0.77, TVP=0.873, TVN=0.67, TFP=0.33, TFN=0.126 et pour le deuxième système nous avons comme résultats : TBC=0.983, TVP=1, TVN=0.967 TFP=0.033, TFN= 0.

Mots clés : Transformée en Ondelettes Continue, SVM, détection des coups de feu.

Abstract

The main objective of this work is to design a gunshot detection module which will be used in audio surveillance systems. We focus in this thesis on the gunshot recognition. Two different systems have been studied. The first system is composed of a single processing stage. It uses the Continuous Wavelet Transform (CWT) for the elaboration of the decision rules. The second system is composed of two processing stages: feature extraction and classification. Feature extraction is based on the CWT for the generation of several energy measurements. The classification stage is based on the Support Vector Machines (SVMs). The sound corpus used in our experiments is extracted from the SOUNDDOG database. The evaluation metrics used in our work are: the True Positive Rate (TPR), the True Negative Rate (TNR), the False Positive Rate (FPR), the False Negative Rate (FNR) and the Classification Accuracy (ACC). The obtained results are satisfactory. For the first system, the results obtained are: ACC = 0.77, TPR = 0.873, TNR = 0.67, FPR=0.33, FNR=0.126 and for the second system, we have obtained the following results: ACC = 0.983, TPR = 1, TNR = 0.967, FPR=0.033, FNR=0.

Key words: Continuous Wavelets Transform, SVM, gunshot detection.

Table des matières

Résumé	(i)
Remerciements	(ii)
Table des matières	(iii)
Liste des tableaux	(iv)
Liste des figures	(v)

Introduction générale:.....	10
-----------------------------	----

Chapitre 1 : Généralités sur les systèmes de surveillance

1.1	Introduction.....	12
1.2	Systèmes de surveillance	12
1.3	Acoustique des coups de feu	13
1.4	Conclusion	17

Chapitre 2 : Détection des coups de feu

2.1	Introduction.....	18
2.2	Détection et reconnaissance des évènements impulsifs.....	18
2.3	Etat de l'art sur les systèmes de détection des coups de feu.....	19
2.4	Détection des coups de feu	20
2.5	Conclusion	37

Chapitre 3 : Expérimentations, résultats et interprétations

3.1	Introduction.....	38
3.2	Logiciels et matériels de développement.....	38
3.3	Métriques d'évaluation	40
3.4	Corpus de tests	42
3.5	Etude expérimentale.....	43
3.6	Application développée	44
3.7	Évaluation des systèmes	45
3.8	Conclusion	48

Chapitre 4 : Conclusions et travaux futures

4.1	Conclusions.....	49
4.2	Travaux futurs.....	50

Annexe	51
---------------------	-----------

Bibliographie.....	55
---------------------------	-----------

Liste d'abréviations et acronymes

CF : *Coups de feu*

SI : *Sons Impulsifs*

SVM : *Support Vector Machines*

TOC : *Transformée en Ondelettes Continue*

ROC : *Receiver Operating Characteristic*

TVP : *Taux de Vrais Positifs*

TVN : *Taux de Vrais Négatifs*

TFP : *Taux de faux positifs*

TFN : *Taux de faux négatifs*

TBC : *Taux de Bonne Classification*

Liste des figures

Figure 1.1 : Système de surveillance de troisiem génération	Erreur ! Signet non défini.
Figure 1.2 : Onde de Friedlander avec temps de montée fini.....	Erreur ! Signet non défini.
Figure 1.3 : Spectre de fréquence d'une onde de Friedlander	Erreur ! Signet non défini.
Figure 2.1 : Système de détection de CF.....	19
Figure 2.2 : Organigramme de système (1).....	22
Figure 2.3 :TOC des signaux à 'M' échelles par le temps 't'	24
Figure 2.4 : Wavelet fonction	25
Figure 2.5 : Organigramme de calcul de la matrice d'énergie	27
Figure 2.6 : Organigramme de système (2).....	29
Figure 2.7 : Recherche d'un hyperplan de séparation optimal.....	31
Figure 2.8 : Hyperplan optimal séparant les observation x1 et x2 de 2 classes Aet B	32
Figure 2.9 : Role de la transformation de l'espace par la fonction φ dans la séparation de données	35
Figure 3.1 : MATLAB R200	39
Figure 3.2 : Logiciel GoldWave.....	39
Figure 3.3 : Principe de la validation croisée	42
Figure 3.4 : Premier application	44
Figure 3.5 : Deuxième application	45
Figure 3.6 : Étude paramétrique	45
Figure 3.7 : Courbe ROC	47

Liste des tableaux

Tableau 3.1 : Matrice de confusion.....	40
Tableau 3.2 : Matrice de confusion de système (1)	46
Tableau 3.3 : Matrice de confusion de système (2)	48

Introduction générale

1. Problématique

La surveillance des activités humaines n'a jamais été aussi omniprésente et massive qu'aujourd'hui. Un nombre important de capteurs sont déployés dans les zones urbaines pour assurer la sécurité des citoyens. En général, les systèmes de surveillance sont basés sur un ou plusieurs capteurs capables d'acquérir des informations provenant du milieu environnant.

Les premiers systèmes de surveillance utilisent seulement des caméras vidéo pour assurer un contrôle permanent des lieux. Le développement de tels systèmes en utilisant seulement les données visuelles entraîne des erreurs considérables. Nous citons : (i) la faible performance des caméras dans des conditions météorologiques défavorables, (ii) leur sensibilité à la commutation légère de lumière, aux reflets et aux ombres [1], (iii) la difficulté de signalisation des événements anormaux qui se produisent en dehors du champ de vision des caméras de surveillance. Pour surmonter ces inconvénients, d'autres types de capteurs ont été utilisés conjointement avec l'information vidéo. Nous citons comme exemples, (i) les capteurs thermiques, (ii) les capteurs infrarouges, (iii) les détecteurs de mouvements, (iv) les capteurs sismiques et (v) les microphones acoustiques. En particulier, l'information audio, captée par les microphones de surveillance, complète efficacement les séquences vidéo vu qu'elle permet de capturer les événements dangereux qui se déroulent en dehors du champ de vision des caméras ou même lorsque les conditions météorologiques sont défavorables.

Dans ce mémoire, nous nous focalisons sur l'utilisation de l'information audio en vue de la détection des coups (CF).

1.2 objectif

Nous contribuons à la mise en œuvre d'un module de reconnaissance des CF en vue de la conception d'un système de surveillance audio. Nous nous intéressons principalement à la différenciation entre les CF et les autres sons impulsifs (SI).

Les systèmes que nous proposons sont implémentés sur un PC ordinaire ayant de bonnes performances. Les critères de choix d'une méthode de reconnaissance efficace sont :

- La complexité algorithmique réduite (faible cout de calcul).
- L'efficacité de détection (taux d'erreurs).
- La robustesse aux bruits de l'environnement.
- Le fonctionnement en temps réel.

1.3 Organisation du mémoire

Le mémoire est organisé comme suit :

Le deuxième chapitre : présent les systèmes de surveillance et les armes à feu. Il présent aussi une étude sur la classification des types d'armes, ainsi qu'une analyse acoustique des CF.

Le troisième chapitre : est consacré à l'approche proposée.

Le quatrième chapitre : contient les résultats expérimentaux.

Le cinquième chapitre : contient une conclusion générale et les travaux futurs.

Chapitre 1 : Généralités sur les systèmes de surveillance et les armes à feu

1.1 Introduction

Nous nous focalisons dans ce chapitre sur les systèmes de surveillance audio. Nous commençons par la présentation des trois générations des systèmes de surveillance. Par la suite. Nous présentons aussi des notions essentielles à la compréhension de l'analyse acoustique des coups de feu (CF). Nous terminerons ce chapitre par une conclusion.

1.2 Systèmes de surveillance

Il existe trois générations différentes :

- La première génération : Ces systèmes sont composés de caméras placées dans des lieux publics pour stocker les activités journalières. Ils sont basés sur des technologies analogiques d'enregistrement. La détection des situations anormales est opérée d'une manière visuelle par les agents de surveillance (en se basant seulement sur les écrans de surveillance). Aucun aspect de l'intelligence artificielle n'est employé [2].
- Les systèmes de surveillance de la deuxième génération « tout numérique » sont définis par une forte intervention avec les technologies de vision par ordinateur (computer vision). Les méthodes de l'intelligence artificielle sont employées pour automatiser le processus d'identification des situations anormales [2].
- La troisième génération : Il s'agit de l'intégration de capteurs complémentaires aux caméras. Différents types de capteurs sont employés, nous citons : les microphones, les capteurs sismiques, les détecteurs de mouvements et les capteurs infra-rouges [2]. Les données issues de ces capteurs sont utilisées pour identifier les anomalies et les situations dangereuses d'une manière automatique par l'intermédiaire des méthodes de l'intelligence artificielle (Figure 2.1).

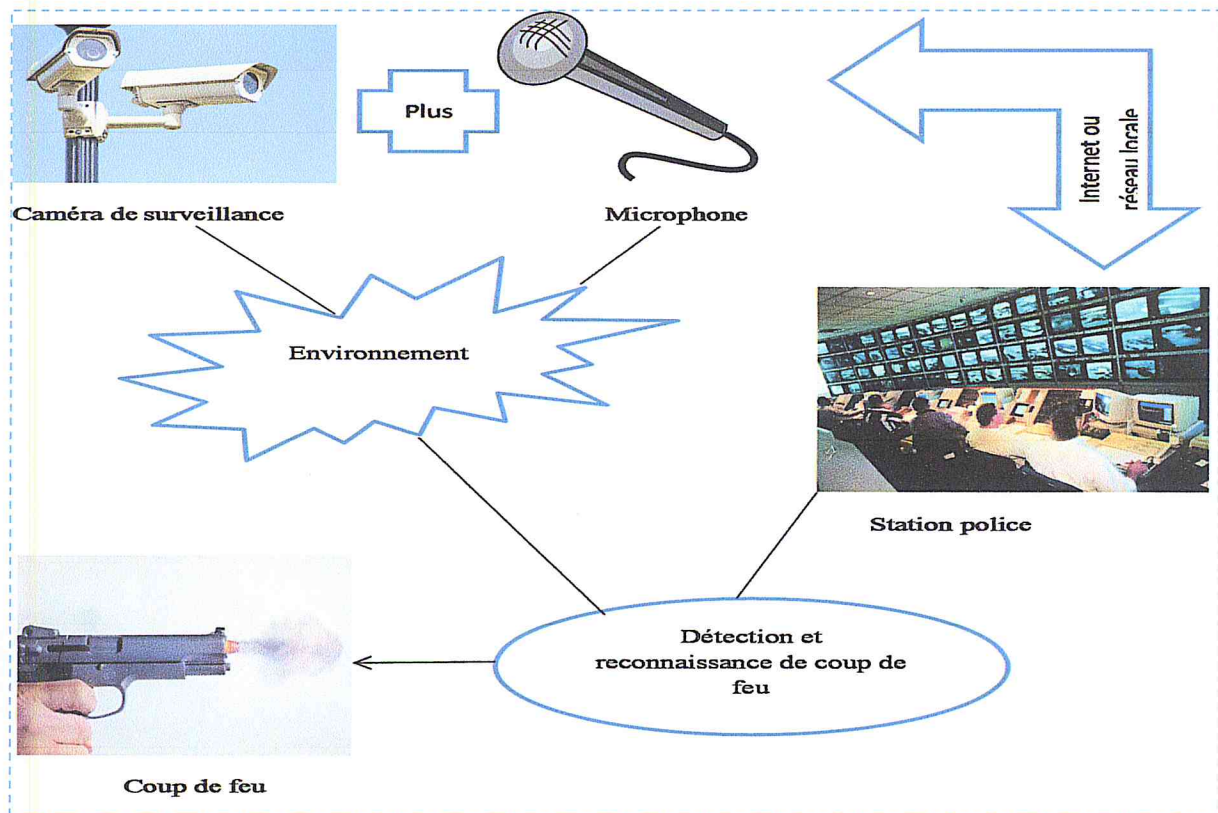


Figure 1.1 : Système de surveillance de troisième génération

1.3 Analyse acoustique des coups de feu

La forme d'onde d'un coup de feu typique est un son impulsif. C'est un signal d'une durée très brève, généralement de moins de 1 seconde. Il est provoqué par une libération soudaine de l'énergie dans l'atmosphère. Ses caractéristiques physiques (acoustiques) sont fortement liées à la géométrie et à l'ampleur de la source [9]. L'analyse des sons est une technique très utilisée pour la détection et la reconnaissance des événements sonores. Les sons des CF peuvent être identifiés par l'analyse de plusieurs paramètres acoustiques. Nous pouvons citer :

- **Puissance et Énergie**

Comme souvent en traitement du signal, on s'intéresse dans le cadre de l'audio à l'énergie. Une mesure standard est la puissance sonore. L'énergie est simplement la somme des amplitudes présentes dans une trame [2].

- **Fréquence**

En physique, la fréquence c'est la hauteur perçue d'un son. Elle est déterminée par le nombre d'impulsions de la perturbation de la pression sonore. L'unité de cette grandeur est le Hertz(Hz) [2].

Il existe trois types différents de sons associés au tir d'une arme : (i) le son d'action mécanique, (ii) l'explosion du Museau et (iii) l'onde de choc.

Les sons mécaniques sont utiles en médecine légale. Ils comprennent trois sons différents dont l'origine est associée respectivement à la gâchette, au marteau et à l'éjection des cartouches usées. Ces sons ne sont pas détectables que si un capteur est placé à proximité de l'arme [10].

3.1.1 L'explosion du museau (Muzel blast)

L'explosion du museau résulte de l'expansion soudaine du gaz suite à l'explosion dans le canon de l'arme à feu. Il rayonne dans toutes les directions se propageant à la vitesse du son. Bien que la pression sonore soit la plus forte dans le sens où le canon de l'arme à feu pointe [10].

L'explosion du museau de l'arme peut être modélisée en monopôle acoustique dans le champ de tir. En outre, l'énergie de l'impulsion sonore est directement proportionnelle au volume de débit de gaz à la source et devient très sensible au bruit de fond et à d'autres sources d'interférences avec une portée croissante. En effet, la pression (énergie) sonore de la source diminue exponentiellement avec le temps et le taux de diminution dépend des caractéristiques d'écoulement de la source [10].

3.1.2 L'onde de choc (Shockwave)

L'onde de choc est dû à la dispersion des molécules d'air provoquée par le mouvement du projectile lors des déplacements à vitesse supersonique et arrive généralement d'abord dans un microphone lorsqu'il est situé dans le domaine de vue de l'onde de choc. Elle dure généralement 0,3 à 0.5 millisecondes. Se caractérise par une augmentation soudaine de la pression suivie d'une baisse approximativement à une valeur presque aussi loin que la montée et puis presque retour instantané à la pression atmosphérique [11].

Pour modéliser les impulsions acoustiques, les ondes Friedlander sont utilisées [10]. L'équation (1.1) correspond à l'expression analytique d'une onde de Friedlander $P(t)$ idéale avec un temps de montée fini [11].

$$P(t) = \left(1 - \frac{t}{c}\right) e^{-t/c}, \quad 0 \leq t < \infty \quad (1.1)$$

Ici le 'c' est la durée d'impulsion mesurée de la surpression au premier passage par zéro ($C=0.65s$), et 't' la durée globale d'explosion [11].

La Figure 1.3 représente une forme d'onde de Friedlander.

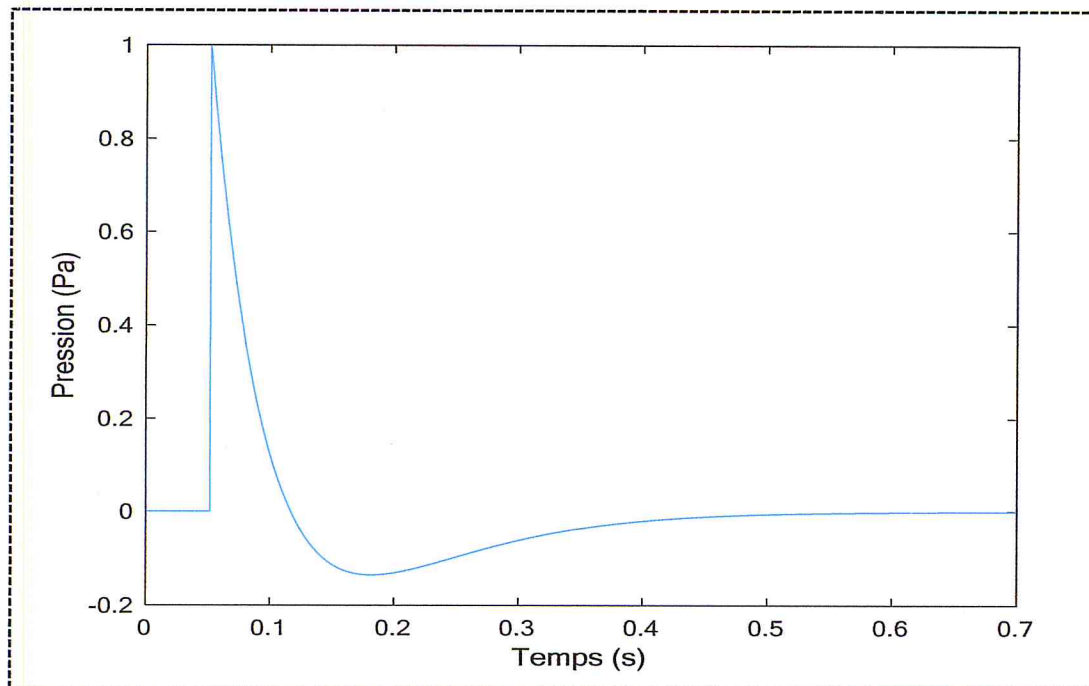


Figure 1.2: Onde de Friedlander avec un temps de montée fini [11].

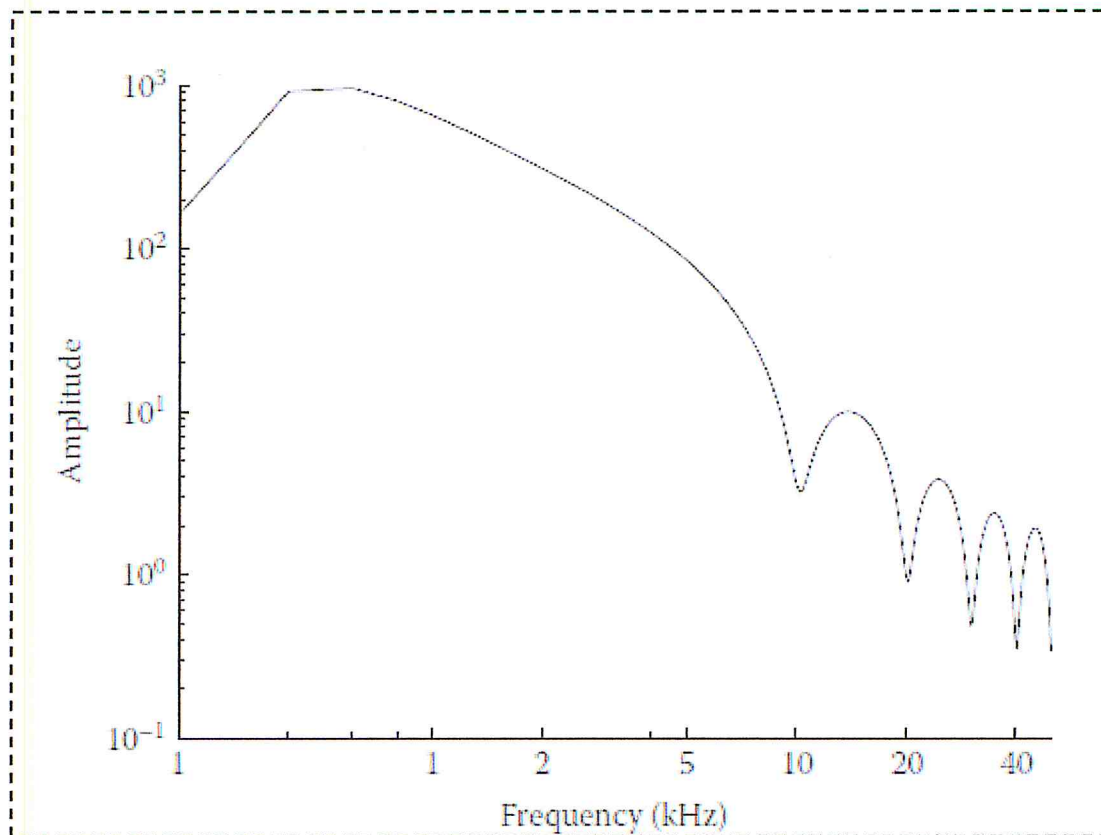


Figure 1.3 : Spectre de fréquence d'une onde de Friedlander [11].

3.1.3 Du l'onde acoustique au signal numérisé

- **Numérisation**

Un système d'acquisition (microphone) permet de transformer une onde acoustique en une onde électrique qui peut être analysée par un ordinateur, ou la numérisation de ce signal est faite pour améliorer la qualité des transformations via deux traitements distincts : l'échantillonnage et la quantification [9].

L'opération d'échantillonnage, consiste à remplacer un signal à temps continu par un signal à temps discontinu. Les échantillonnages sont espacés d'une durée définie par la fréquence d'échantillonnage ; à la suite à partir de ces échantillons nous pouvons calculer des grandeurs physiques (paramètres acoustiques) nommées descripteurs comme les paramètres cités précédemment [9].

La quantification, est la représentation numérique de la valeur de chaque échantillon dans l'espace des valeurs possibles [9].

La signature acoustique des SI est souvent modélisée par l'onde de Friedlander. Selon le spectre de fréquences d'une onde de Friedlander, l'énergie d'un CF inversement proportionnelle est l'augmentation de ses fréquences. Selon la Figure 2.8, les fréquences entre 500 Hz et 800Hz possèdent une énergie élevée par rapport aux autres fréquences (zone d'intérêt).

1.4 Conclusion

Dans ce chapitre, nous nous sommes focalisés sur les systèmes de surveillances et les armes à feu. Nous avons présenté en premier lieu les différents types des systèmes de surveillance. Par la suite, nous avons exposé les armes à feu ainsi que leur signature acoustique. Dans le prochain chapitre, nous exposerons les méthodes de détection des CF.

Chapitre 2 : Détection des coups de feu

2.1 Introduction

Dans ce chapitre, nous présentons quatre sections importantes : (i) la détection et reconnaissance des événements impulsifs, (ii) état de l'art sur les méthodes de reconnaissance des CF, (iii) la description des systèmes proposées. Nous terminons ce chapitre par une conclusion.

2.2. Détection et reconnaissance des événements impulsifs

Un système de détection des coups (Figure 3.1) est composé de deux modules essentiels :

- ✓ Le premier module concerne la détection des SI. Il s'agit de l'opération de différenciation entre les sons impulsifs (SI) et les autres sons de l'environnement. Nous citons quelques exemples de SI : les sons de claquement de portes, des cris humains, des sirènes, verres cassés et les CF. Les autres sons de l'environnement peuvent être de nature variable : les bruits de l'environnement, les discours, les sons de la nature et autres. Ce module ne fournit pas des informations sur le type de SI, mais il fournit une information qu'un SI s'est produit.
- ✓ Le deuxième module concerne la reconnaissance des CF. Il vise à reconnaître le type exact des SI produit. Donc il s'agit de la différenciation entre les CF (un cas particulier des SI), et les autres SI (non CF).

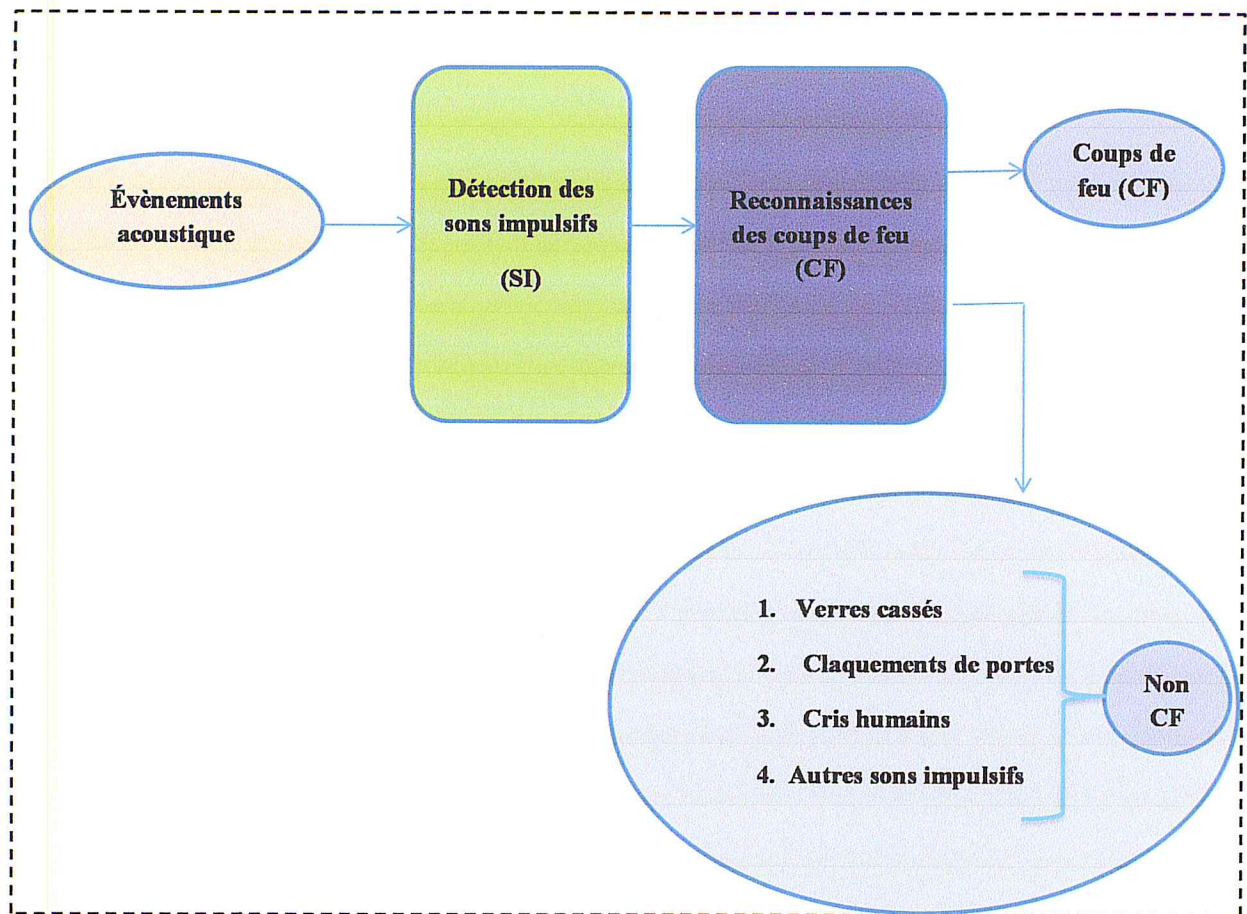


Figure 2.1 : Système de détection des coups de feu.

2.3 État de l'art sur les systèmes de détection des coups de feu

La détection des CF est un domaine de recherche actif en traitement du signal [12] [13]. Dans ce genre d'applications, les algorithmes et les méthodes de détection doivent être simples et non-complexes. La complexité des méthodes choisies est un critère de choix pour la détection des CF en temps réel.

De nombreuses solutions dans l'état de l'art exploitent des méthodes de reconnaissance de formes pour détecter les CF.

Nous présentons par la suite deux solutions de détection de CF reportés dans la littérature :

- Talal et al. [12] ont proposé une approche de détection de CF basée à deux étages de traitement, la détection des évènements acoustiques et la reconnaissance des CF. L'étape de détection utilise un formalisme probabilistique à base d'une distribution Chi-square. La reconnaissance est basée sur les modèles SVMs (Support Vector Machines). Cette méthode a donné de bonnes performances de reconnaissance.

- A. Rabaoui et al. [13] ont exploité les modèles SVMs à une seule classe pour détecter et classifier des évènements audio. Plusieurs catégories de sons ont été utilisées. Nous citons : les CF, les cris humains, les voix des enfants, les explosions et autres. L'étape de détection des sons est basée sur une approche combinée entre une famille de fonctions dites « exponentielles » avec un modèle SVM à une seule classe. Cette solution vise à approximer le rapport de vraisemblance généralisé. L'étape de classification des sons est basée aussi sur un formalisme à base des SVMs. Les résultats obtenus sont bien adaptés aux tâches de reconnaissance d'évènements.

2.4 Détection des coups de feu

Notre objectif dans le cadre de ce projet de fin d'étude est de mettre en œuvre un détecteur des CF en temps réel. Les critères de choix d'une méthode de détection efficace sont :

- La complexité algorithmique et le cout de calcul,
- L'efficacité de détection (taux d'erreurs),
- La robustesse face aux bruits de l'environnement,
- Le fonctionnement en temps réel.

Parmi les méthodes de reconnaissance qui existent dans la littérature, nous avons trouvé que les modèles SVMs répondent mieux aux critères ciblés ci-dessus. Les SVMs sont des modèles qui ont prouvé leurs efficacités dans plusieurs domaines d'applications (reconnaissance du visage, diagnostic des maladies CT-Scan, reconnaissance des galaxies et autres). De plus, l'existence de la bibliothèque LIBSVM (C++, JAVA, Python, MATLAB et autres) facilite l'utilisation de ce formalisme mathématique.

Nous avons proposé deux systèmes différents pour la reconnaissance des CF. Le premier système est composé d'un seul étage de traitement. Elle consiste en l'utilisation de la transformée en ondelettes continue (TOC) en vue de l'élaboration des règles de décision. Le deuxième système est composé de deux étages de traitement, l'extraction d'attributs et la classification. L'extraction d'attributs est basée aussi sur la TOC pour la génération de plusieurs mesures énergétiques. L'étage de classification est basé sur les Support Vector Machines (SVMs).

Les deux systèmes exploitent des données acoustiques de CF pour la mise en œuvre des procédés de décisions.

- **Données acoustiques**

Nous avons utilisé trois types d'armes à feu pour la mise en œuvre des modèles mathématiques de décision : (i) Revolvers (ii) Pistolets et (iii) Carabines. Les CF représentent la classe positive. La classe négative est représentée par d'autres évènements dangereux (impulsifs) : (i) aboiement de chiens, (ii) cries des humains et (iii) claquement de porte.

2.4.1 Premier système

Nous avons proposé une approche de décision à base de règles statistiques pour identifier les CF. Cette approche n'est pas basée sur des machines d'apprentissage (classifieurs) (Figure 3.2) pour la composition du système de reconnaissance.

Nous avons utilisé la transformée en ondelettes continue (TOC) pour l'élaboration de ces règles. Néanmoins, l'exactitude de détection dépend d'une seule contrainte : La sélection des paramètres algorithmiques. Dans cette section, nous donnons une description détaillée sur le système proposé. Nous présentons les différents paramètres algorithmiques ainsi que les principales étapes de détection.

2.4.1.1 Description

Le système est schématisé sur la Figure 2.2. Il est composé de trois étapes principales :

1^{er}étape : calcul de la TOC (matrice des coefficients)

Nous avons choisi la transformation en ondelettes continue (TOC) comme méthode pour l'analyse temps-fréquence [14].

Au moyen d'une fonction spécifique Ψ , l'analyse est réalisée. Pour sélectionner la partie du signal a traitée, cette ondelette est positionnée dans le domaine temporel. Puis elle est dilatée ou contractée selon la valeur d'échelle pendant l'analyse [14]. La TOC est capable de fournir une infinité d'échelles. Elle permet la décomposition temporelle du signal par contraction et dilatation corrélée à l'échelle. La TOC est donné comme suit :

$$s(a, b) = |a|^{-\frac{1}{2}} \int s(t) * \Psi\left(\frac{t-b}{a}\right) dt \quad (2.1)$$

$s(t)$: le signal original,

a : paramètre d'échelle $a \neq 0$,

b : paramètre de translation,

Ψ : Ondelette analysante.

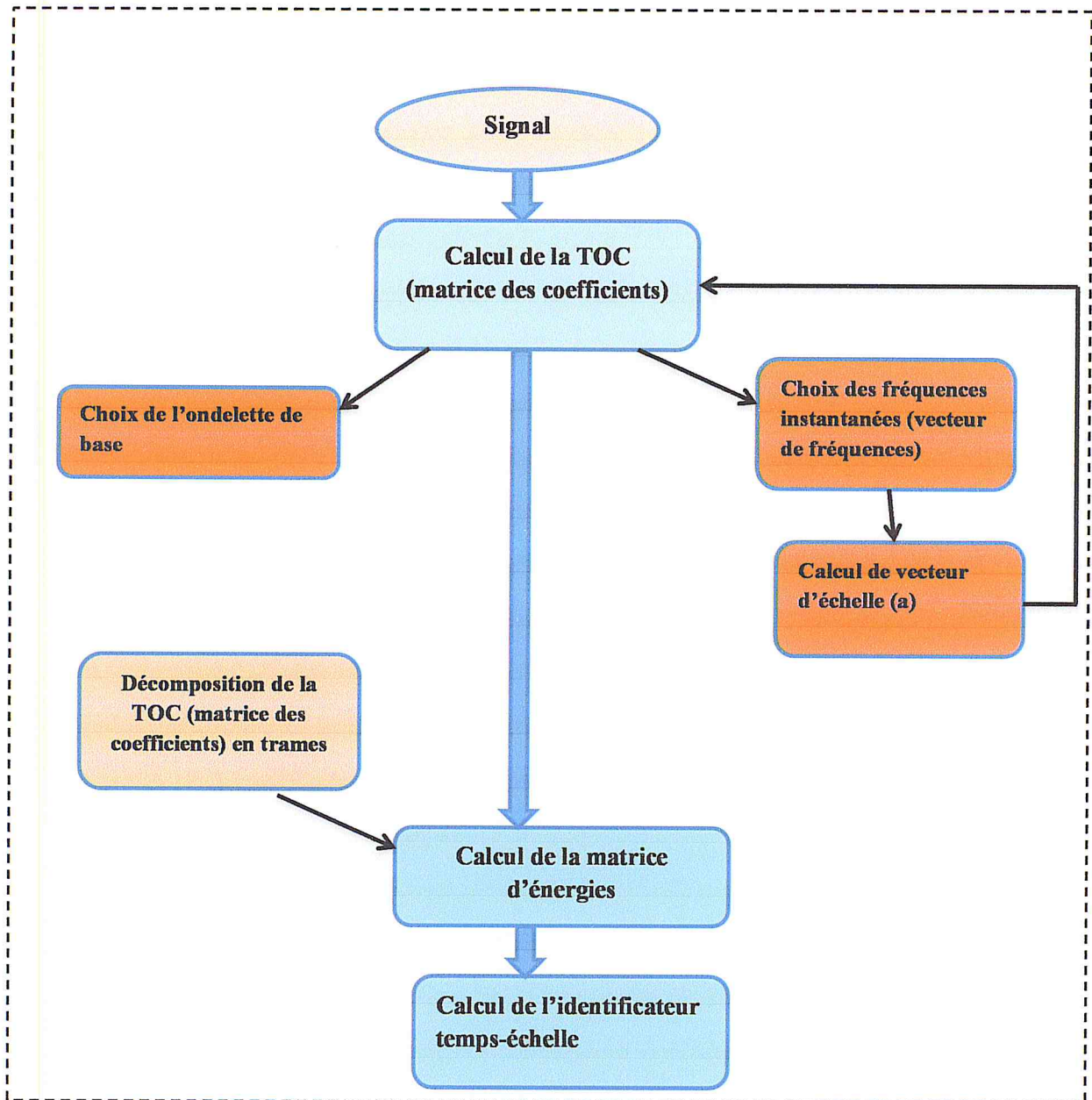


Figure 2.2 : Organigramme du système (1).

La relation entre l'échelle « a » et la fréquence instantanée « Freq » peut être approximée par la formule suivante :

$$a = \frac{F_{\text{centre}}}{\left(\text{Freq} * \left(\frac{1}{f_e} \right) \right)} \quad (2.2)$$

F_{centre} : est la fréquence centrale de l'ondelette. Sa valeur est dépendante du type d'ondelettes choisi (Haar, Daubechies, Coiflets, Symlets, Morlet, Mexican Hat et Meyer).

f_e : est la fréquence d'échantillonnage du signal $s(t)$.

✓ **Choix des fréquences instantanées (vecteur de fréquences)**

Selon l'analyse acoustique des CF, nous avons remarqué que le spectre des fréquences de ces événements fournit des informations pertinentes sur leurs natures acoustiques. A cet effet, la bande fréquentielle utilisée pour le calcul de la TOC ainsi que l'ensemble des fréquences instantanées doivent être choisies minutieusement.

Nous avons choisi dans notre cas d'exploiter l'échelle fréquentielle de Mel [15], pour atteindre cet objectif. En effet, l'échelle de Mel est une échelle psycho acoustique des fréquences dont l'unité est le **Mel**. Il est relié au hertz (**Hz**) par une relation basée sur une étude expérimentale, sur l'audition humaine.

$$Mel = 2595 * \log_{10} \left(1 + \left(\frac{f}{700} \right) \right) \quad (2.3)$$

Où :

Mel : vecteur de l'échelle fréquentielle de Mel (unité **Mel**).

f : est le vecteur de l'échelle fréquentielle du système international (unité **Hz**).

La transformation inverse est donnée par la formule suivante :

$$f = 700 * \left(\left(10^{\left(\frac{Mel}{2595} \right)} - 1 \right) \right) \quad (2.4)$$

Une fois le vecteur de fréquences $V(f)$ est sélectionné, nous pouvons déduire facilement l'échelle « a ». Il sera par la suite utilisé pour calculer les coefficients de la TOC.

Nous avons suivi ces étapes pour calculer $V(f)$:

1. Définir une bande fréquentielle B_F en unité de **Mel**. Elle doit varier entre une fréquence basse (F_{Mel_basse}) et une fréquence haute (F_{Mel_haute}) tel que $B_F = [F_{Mel_basse}, F_{Mel_haute}]$ **Mel**.
2. Définir le nombre des fréquences Mel instantanées (M).
3. Générer linéairement un vecteur fréquentiel $V_{Mel}(f)$ contenant M fréquences (en unité de **Mel**) variant entre F_{Mel_basse} et F_{Mel_haute} .
4. Convertir la suite $V_{Mel}(f)$ (initialement en unité de **Mel**) en une nouvelle suite $V(f)$ en unité de **Hz**.

✓ **Calcul de vecteur d'échelle « a »**

Le vecteur d'échelle « a » est calculé à partir du vecteur des fréquences $V(f)$ en utilisant l'équation 3.2 où $Freq$ est remplacée par $V(f)$. La taille du vecteur représente le nombre d'échelle « M » (Figure 3.3).

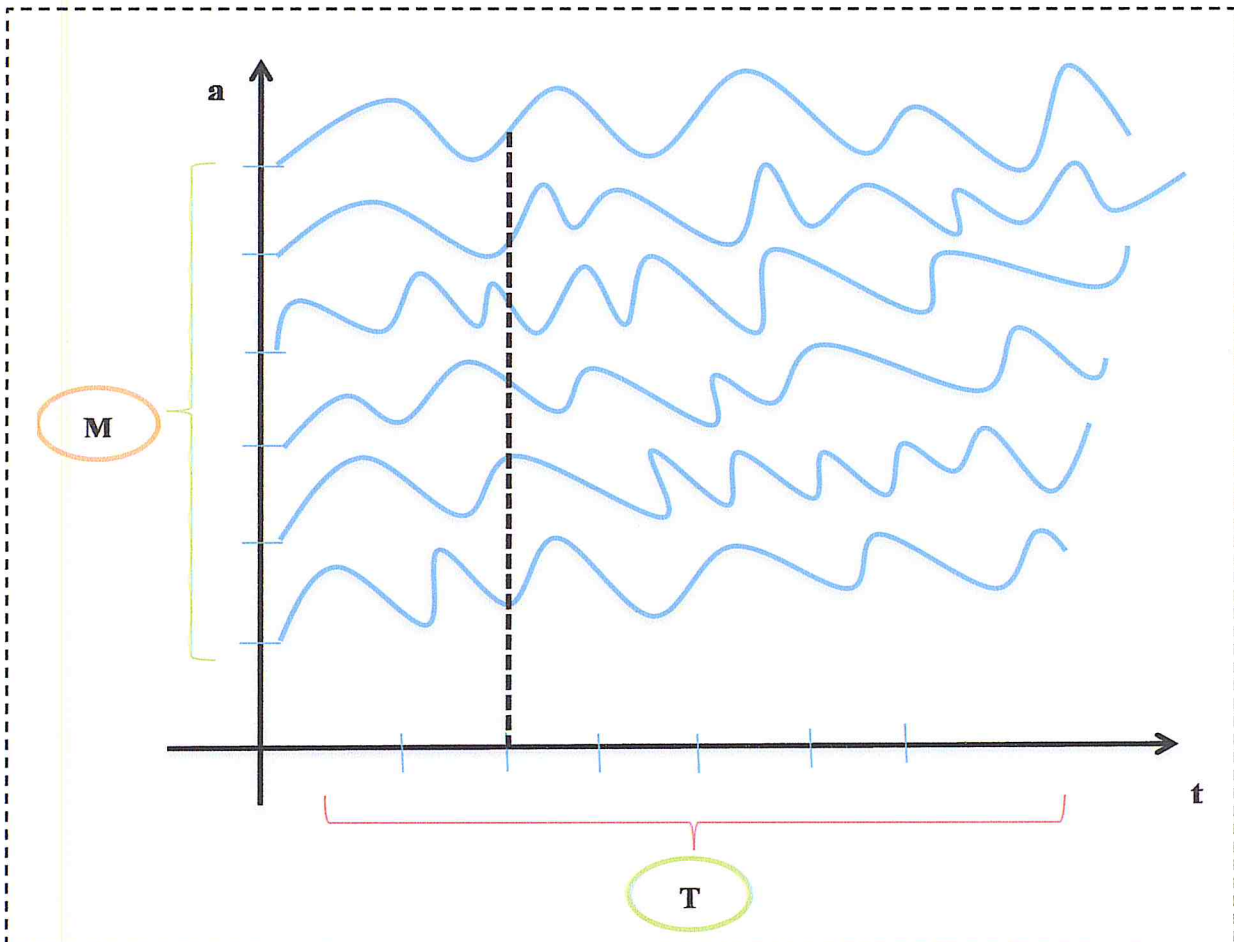


Figure 2.3 : TOC des signaux à 'M' échelles par le temps 't'.

Ou :

- T : représente la longueur du signal $s(t)$.
- M : le nombre d'échelles.

✓ **Choix de l'ondelette de base**

Pour réaliser une analyse temps-échelle, il faut d'abord choisir la fonction analysante de base (ondelettes mère). Il existe plusieurs familles d'ondelettes, la plus utilisée est l'ondelette de Daubechis (Db) [16], nous avons utilisé Daubechis avec un ordre 15 (Db15), pour une bonne analyse temps-fréquences (échelles) (Figure 2.4).

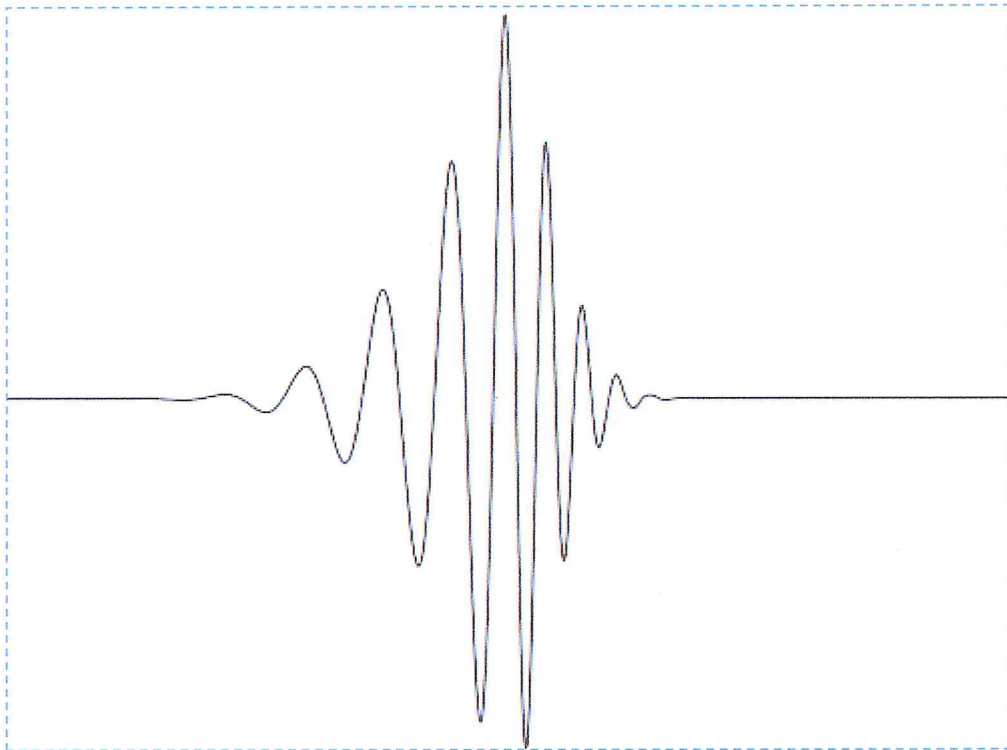


Figure 2.4 : Fonction d'ondelette.

Il y a cinq étapes simples pour la transformée en ondelette continue [14].

- ✓ Choisir le type de l'ondelette qui peut être positionnée dans le domaine temporel.
- ✓ Multiplier le signal par l'ondelette de fréquence $1/a$ décalée de b par rapport à l'origine des temps. Puis calculer le coefficient d'ondelette S pour chaque valeur de a et de b par la somme algébrique du produit obtenue. Donc le résultat dépend à l'ondelette choisie.
- ✓ Translater l'ondelette à gauche et répéter les étapes 1 et 2 jusqu'à la fin du signal.
- ✓ Passer à la prochaine échelle et répéter les étapes de 1 à 3.
- ✓ Répéter les étapes de 1 à 4 pour toutes les échelles.

Après le déroulement de toutes les étapes, nous obtenons une matrice de coefficients MAT de taille $M \times T$. M et T représentent respectivement le nombre d'échelles et la longueur du signal $s(t)$.

2^{ème} étape : calcul de la matrice d'énergie E

Après le calcul de la matrice des coefficients MAT, nous la découpons verticalement en N' trames. La valeur de N est un paramètre à optimiser. Nous présentons une approche d'optimisation des paramètres de la méthode dans le chapitre des résultats.

Par la suite, et pour chaque trame, nous calculons la matrice des énergies E(i,j) à partir d'un ensemble de signaux à court terme $x_{i,j}$ qui sont extraits de la matrice MAT. Nous obtenons ainsi « M » composantes pour chaque trame. Le calcul des mesures énergétiques est basé sur la formule suivante :

$$E(i, j) = \frac{1}{Nk} \sum_{k=0}^{Nk-1} x_{i,j}(k)^2 \quad (2.5)$$

Où :

i : est l'indice des échelles. Il varie de 1 à M,

j : est l'indice des trames. Il varie de 1 à N,

k : est l'indice des échantillons. Il est réservé aux signaux à court terme de chaque trame,

Nk : est le nombre d'échantillons des signaux à court terme, $Nk = \left[\frac{T}{N} \right]^1$

$x_{i,j}$: est le i^{ème} signal à court terme de la j^{ème} trame,

La matrice d'énergies E est de taille N×M. Le processus de calcul de E est récapitulé sur la (figure 2.5).

¹ [*] représente la partie entière de la fraction

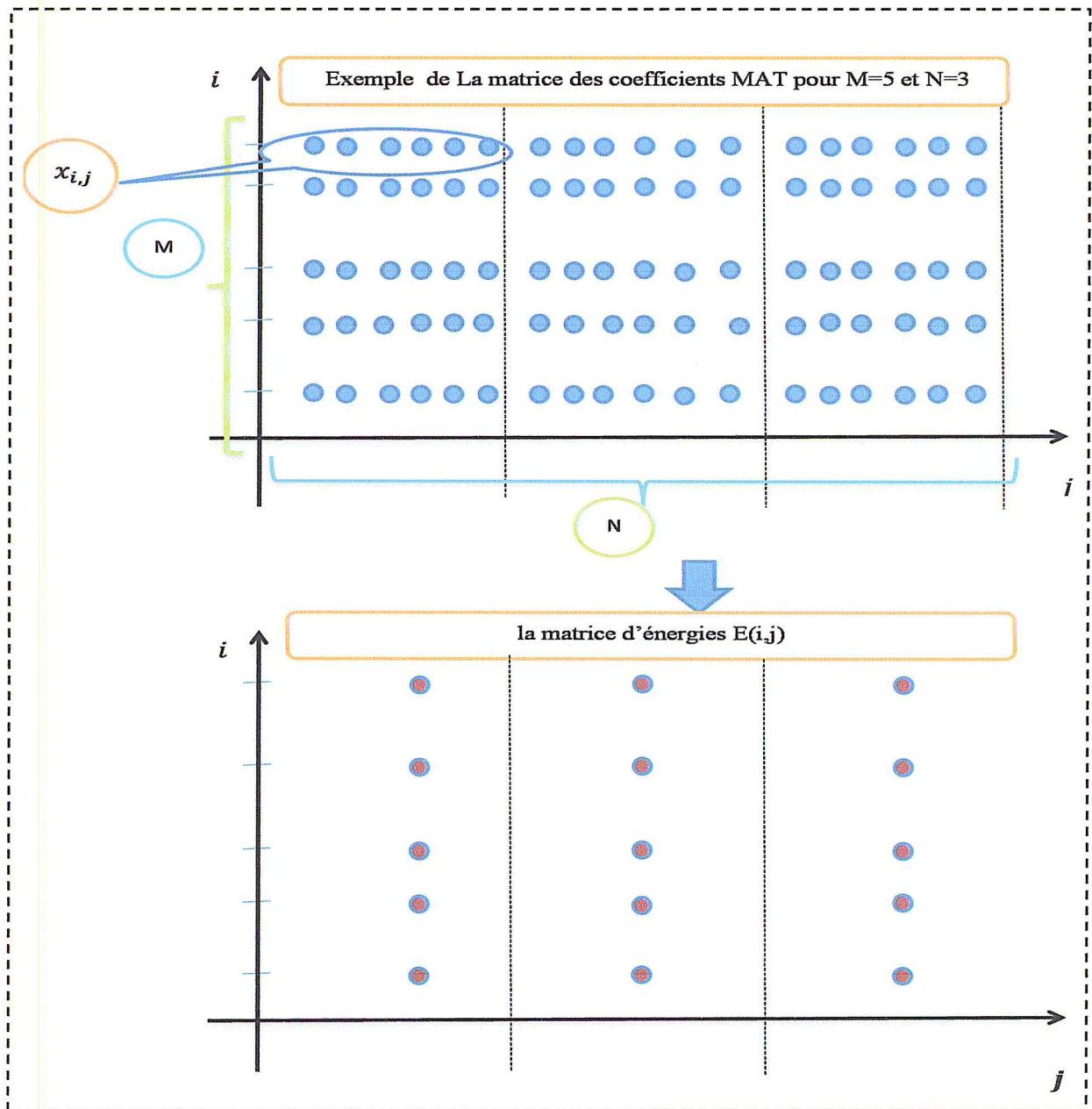


Figure 2.5 : Organigramme de calcul de la matrice d'énergie

3^{me} étape : Calcul de l'identificateur temps-échelle

Dans cette étape, nous calculons une entité statistique qui représente la valeur à énergie maximale d'un signal acoustique : IDmaxechelle. La règle de décision, concernant l'appartenance d'un évènement acoustique à la classe positive (CF) ou la classe négative (autres sons impulsifs), sera élaborée à base de cet identificateur (Figure 2.6).

Trois étapes importantes doivent être exécutées pour obtenir IDmaxechelle :

1. Pour chaque colonne de $E(i,j)$, nous calculons la matrice des pourcentages $M_p(i,j)$ comme suit :

$$M_p(i,j) = \frac{100 * E(i,j)}{\sum_{i=1}^M E(i,j)} \quad (2.6)$$
2. Ensuite, nous cherchons l'élément maximal de la matrice $M_p(i,j)$. L'indice 'i' de cet élément correspond à $ID_{maxechelle}$.
3. Si $ID_{maxechelle} = 1$, l'événement impulsif est classé CF.

3.4.1.2 Positionnement du problème

La complexité de la méthode proposée ci-dessus est liée aux paramètres suivants :

- ✓ La fréquence basse « F_{Mel_basse} »,
- ✓ La haute fréquence « F_{Mel_haute} »,
- ✓ Le nombre de trame « N »,
- ✓ Le nombre d'échelles « M ».

Vu que le système proposé est dédié à la reconnaissance en temps réel des CF, nous sommes obligés d'optimiser ces paramètres en vue de minimiser la complexité algorithmique de la solution. Nous avons opté pour une procédure expérimentale d'optimisation à base d'une méthode appelée **Grid search** est présentée dans le chapitre prochain.

La question importante qui se pose lors de l'implémentation de la méthode est la suivante :

- Es ce que le choix d'un identificateur ($ID_{maxechelle}$) est suffisant pour reconnaître toutes les catégories des CF ?

La réponse à cette question sera faite après les résultats expérimentaux.

2.4.2 Deuxième système

Le deuxième système que nous proposons est basée sur une approche à deux étages de traitement : (i) l'extraction d'attributs et (ii) la classification. Ce système est représenté par le schéma suivant :

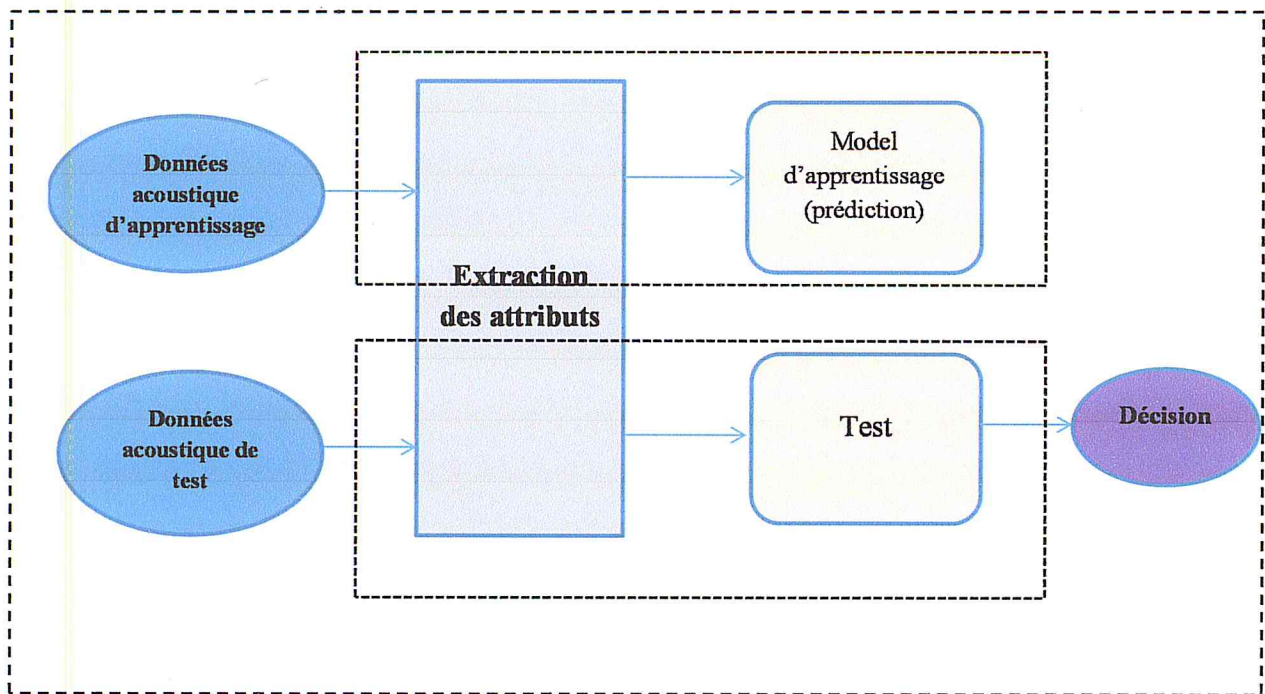


Figure 2.6 : Organigramme de système (2)

2.4.2.1 Description

L'extraction d'attributs est basée sur la TOC. La classification est basée sur les SVMs.

Dans cette section, nous donnons une description détaillée de la méthode de détection. Nous présentons les différents paramètres de cette méthode ainsi que les principales étapes de détection.

1^{er} étape : Extraction d'attributs

Cette étape consiste à extraire le vecteur d'attributs, qui est représenté dans notre cas par des mesures énergétiques d'un signal acoustique. Il s'agit de transformer la matrice des énergies $E(i,j)$, qui a été utilisée dans le premier système, sous une forme vectorielle en vue de tenir en compte de l'évolution temporelle du signal $s(t)$. Le vecteur résultant est nommé V_{eng} .

La dimension de la matrice $E(i,j)$ est : $M*N$.

Le vecteur d'énergies est obtenu en concaténant les N colonnes de la matrice $E(i,j)$:

$$V_{eng} = [E(:,1), E(:,2), \dots, E(:,N)]$$

$V_{eng}(l)$ est d'une longueur de $M*N$. $l=1 : M*N$.

Il faut souligner que l'optimisation des autres paramètres nécessaires à l'extraction d'attributs doivent être aussi optimisés pour avoir de bonnes performances de reconnaissance. Il s'agit de :

- ✓ La fréquence basse « F_{Mel_basse} »,

- ✓ La haute fréquence « F_{Mel_haute} »,
- ✓ Le nombre de trame « N »,
- ✓ Le nombre d'échelles « M ».

2^{ème} étape : Classification

Pour reconnaître si un évènement impulsif est un CF ou non, nous avons utilisé un SVM à 2 classes. Avec ce concept, nous pouvons classer un ensemble de points de données en deux classes en utilisant une fonction de décision Gaussienne (**RBF : Radial Basis Function**).

A. Machines à vecteurs de supports

Les machines à vecteurs de supports (Support Vectors Machines SVMs), est une méthode de classification par apprentissage. Elle fut introduite par Vapnik en 1995. Elle a été tout d'abord exposée dans le cas des problèmes linéairement séparables. Elle est basée sur l'utilisation d'une fonction de décision dite noyau (kernel), qui n'utilise qu'un sous ensemble de la base d'apprentissage. Les constituants de ce sous-ensemble sont les vecteurs de support.

D'une part, SVM permet de résoudre les problèmes linéairement séparables et d'avoir un classifieur linéaire dans un espace approprié. D'autre part, elle peut résoudre aussi des problèmes non linéairement séparables, et sélectionne un classifieur par une projection tacite dans un espace de dimension supérieur avec un séparateur optimale des données [2] [1].

○ Principe de fonctionnement

● Problème :

Comme tous les problèmes d'apprentissage, on considère une variable Y à prédire, et pour la simplification est supposé à valeurs dans $\{-1,1\}$, et soit $X = \{x_1, \dots, x_n\}$, les variables prédictives,

On considère aussi l'ensemble d'apprentissage présenté par des couples (vecteur d'attributs, label). Le vecteur d'attributs représente les variables prédictives. Le label désigne les étiquettes des variables à prédire [24].

On note

$$E = \{(x_1, y_1) \dots (x_n, y_n)\} \quad (2.7)$$

$$x_i \in X \text{ et } y_i \in Y$$

E : est l'ensemble d'apprentissage.

(x_i, y_i) : couple (vecteur d'attributs , label).

Dans ce cas, le problème se pose comme la recherche d'une frontière de décision qui maximise la distance entre les vecteurs de support, et qui sépare les observations x_i des deux classes, dans un espace H , par une fonction non-linéaire $\varphi : R^n \rightarrow H$.

Cette fonction permet de projeter les variables prédictives (observations), dans ce nouvel espace de travail H . La frontière entre les deux classes est matérialisée par un hyperplan optimal parmi les séparateurs valides (Figure 2.7). On cherche alors à trouver le compromis concernant cette frontière [24].

La classe y d'un nouvel exemple x est définie par une fonction f dont le signe désigne la fonction d'appartenance à une classe donnée.

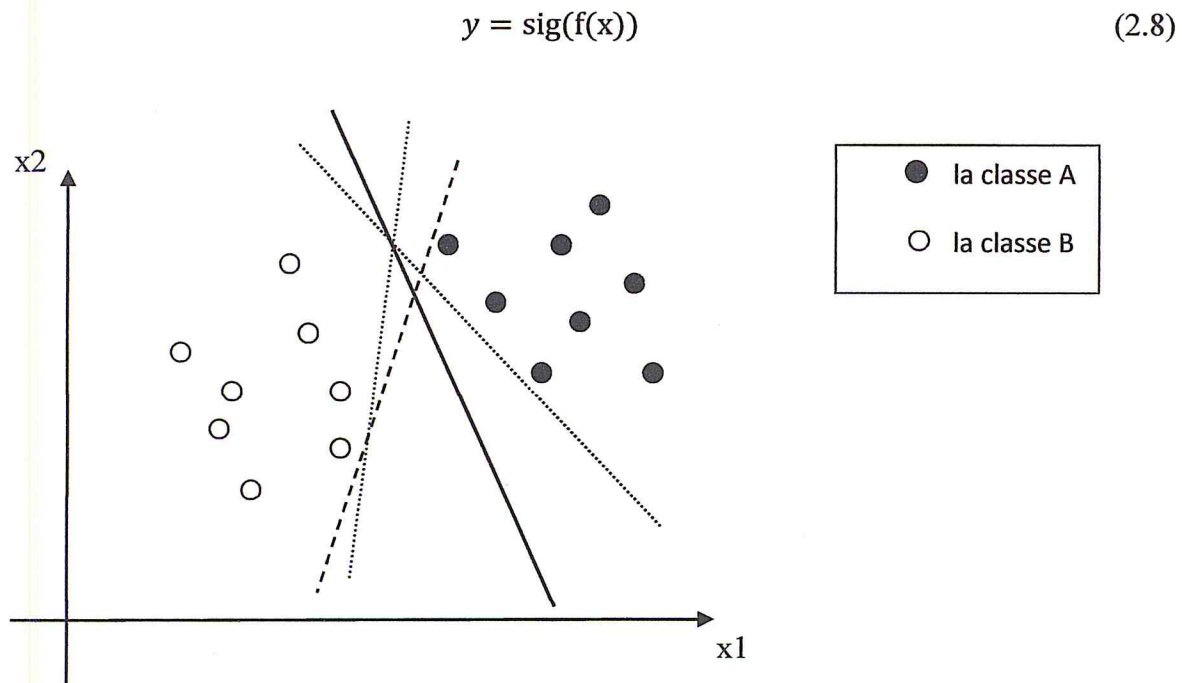


Figure 2.7: Recherche d'un hyperplan de séparation optimal.

a) Hyperplan

Selon l'information citée par Sébastien_Lecomte [2], « Le terme hyperplan est une généralisation en dimension 'd' quelconque de la notion de plan, dans l'espace ou encore de droite dans le plan. Autrement dit, cela correspond à une frontière réalisant une partition complète, et exclusive de cet (hyper-)espace de dimension 'd'. » [2].

Autrement un hyperplan est la frontière qui sépare les données d'apprentissage des deux classes. Il faut souligner qu'il existe plusieurs hyperplans qui sont valides.

b) Espace de re-description

Consiste à transformer les valeurs de X par une fonction φ dans un espace H de plus grande dimension. L'objectif de cette transformation est de traiter un problème non-linéaire, et d'avoir la possibilité de la séparation linéaire des données dans ce nouvel espace.

Cette transformation est fondamentale dans le principe des SVMs [2].

B. Séparateur linéaire

a) Hyperplan séparateur

Selon les propriétés des SVMs, et vu qu'il est possible de trouver plusieurs séparateurs de données, l'hyperplan optimal est sélectionné parmi tous les hyperplans valides dans le cas où la séparation est possible. Pour cela on peut dire que c'est un hyperplan séparateur [25].

Les points les plus proches, qui seuls sont utilisés pour la détermination de l'hyperplan, sont appelés vecteurs de supports. La distance entre les vecteurs de supports et l'hyperplan optimale doit être maximale. Elle est nommée marge maximale (Figure 2.8).

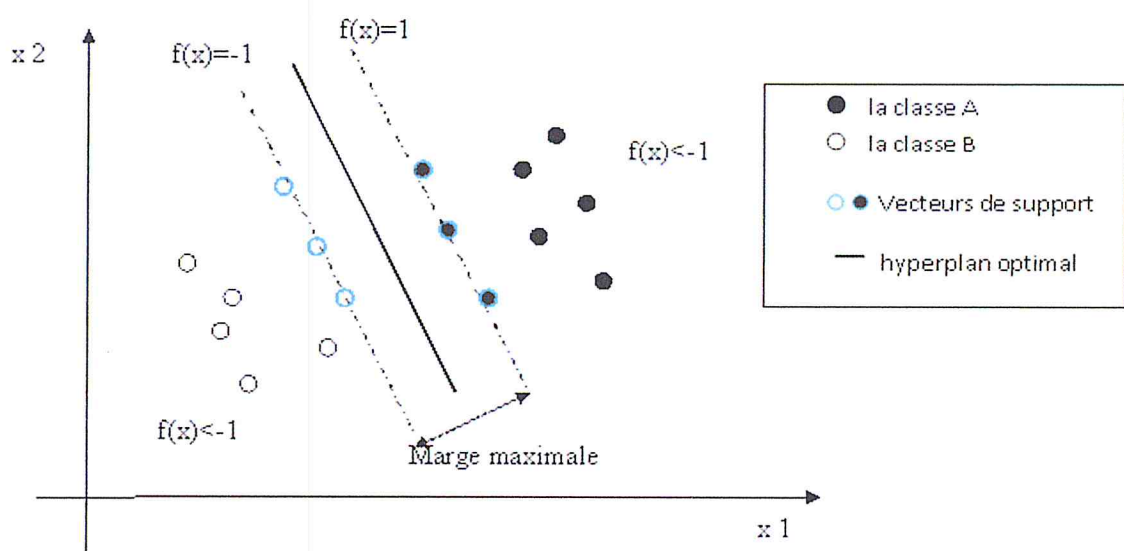


Figure 2.8 : Hyperplan optimal séparant les observations x_1 et x_2 de deux classes A et B

Dans le cas linéaire, un hyperplan muni d'un produit scalaire est défini par l'équation :

$$\langle w, x \rangle + b = 0 \quad (2.9)$$

* b est le biais et $\frac{b}{\|w\|}$ est la distance de l'hyperplan à l'origine.

*w est la normale à cet hyperplan tandis que la signe de la fonction $f(x) = \langle w, x \rangle + b$ indique de quel coté se trouve x à prédire. Il sera bien classé si :

$$yf(x) \geq 1 \quad (2.10)$$

Un plan (w,b) est un séparateur si

$$y_i f(x_i) \geq 1 \quad \forall i \in \{1, \dots, n\} \quad (2.11)$$

Nous avons la distance di d'une observation x au plan(w, b) [2], et nous appelons marge la valeur M tel que $d \geq M$. La distance est donnée par :

$$d(x_i) = \frac{y_i f(x_i)}{\|w\|} \geq M \quad (2.12)$$

D'après (2.10) et (2.12) :

$$\|w\| M = 1 \quad (2.13)$$

On obtient la marge $M = \frac{1}{\|w\|}$ qui doit être maximale (par la minimisation de $\|w\|$)

Pour choisir l'hyperplan séparateur optimal, on doit résoudre la formulation primale [25]:

$$\min \frac{1}{2} \|w\|^2 \quad (2.14)$$

sous les contraintes $y_i(w^T x_i + b) \geq 1, i = 1, \dots, m$

On passe du problème primal au problème dual, en introduisant des multiplicateurs de Lagrange pour chaque contrainte dans le cas où la dimension de l'espace des observations X dépassant quelques centaines [25].

$$\text{Max} \left\{ \sum_{i=1}^m \alpha_i - \left(\frac{1}{2}\right) \sum_{i,j=1}^m \alpha_i \alpha_j y_i y_j \langle x_i, x_j \rangle \right\}$$

$$\alpha_i \geq 0, \quad i = 1, \dots, m \quad (2.15)$$

$$\sum_{i,j=1}^m \alpha_i y_i = 0$$

L'hyperplan solution correspondant peut alors s'écrire :

$$f(x) = (w^* \cdot x) + b^* = \sum_{i,j=1}^m \alpha_i^* y_i \langle x_i, x_j \rangle + b^* \quad (2.16)$$

Où α_i sont les solutions de l'équation (3) et b^* est obtenue en utilisant n'importe quel exemple (x_i, y_i) dans l'équation :

$$\alpha_i [y_i ((x_i \cdot w^*) + b) - 1] \quad (2.17)$$

b) Cas non séparable

Lorsque les exemples ne sont pas séparables par un hyperplan, on considère la fonction f et on introduit des variables de relâchement, avec une précision ε_i pour rendre faisables les contraintes du problème d'optimisation [2] :

$$y_i f(x_i) \geq 1 - \varepsilon_i \quad (2.18)$$

Alors, pour chaque exemple (x_i, y_i) , il existe une certaine erreur $\varepsilon_i > 0$ vérifiant la nouvelle contrainte. Le problème primal devient [2]:

$$\min \frac{1}{2} \|w\|^2 + \frac{1}{Nu * m} \sum_{i=1}^m \varepsilon_i$$

$$\text{sous les contraintes } y_i (w^T x_i + b) \geq 1 - \varepsilon_i, i = 1, \dots, m \quad (2.19)$$

$$\varepsilon_i \geq 0$$

Remarques [2]

Dans cette formule, c'est le paramètre Nu qui caractérise la solution.

1. Il établit une limite supérieure sur la fraction des valeurs aberrantes (exemples d'apprentissage considérés comme autres classe).

2. C'est une limite inférieure sur le nombre d'exemples d'apprentissage utilisés comme SVMs.

Le problème dans le cas non séparable se met sous la même forme duale, que dans le cas séparable à une différence près : les coefficients α_i sont tous bornés par la constante de contrôle de la pénalisation Nu [2] :

$$\alpha_i \geq 0, \quad i = 1, \dots, m \quad \Rightarrow \quad 0 \leq \alpha_i \leq \frac{1}{\text{Nu} \cdot m} \quad i = 1, \dots, m \quad (2.20)$$

C. Séparateur non-linéaire

a) Noyau

Pour obtenir un problème linéaire simple à partir d'un cas non linéaire, la projection des données est faite dans un espace de dimension plus grand. On peut alors avoir une possibilité très élevée de trouver un hyperplan séparateur des observations.

Pour cela en prend un exemple :

Soit la projection φ définie par

$$\begin{aligned} \varphi : R^2 &\rightarrow R^3 \\ \varphi(x_1, x_2) &= (x_1^2, \sqrt{2} x_1 x_2, x_2^2) \end{aligned} \quad (2.21)$$

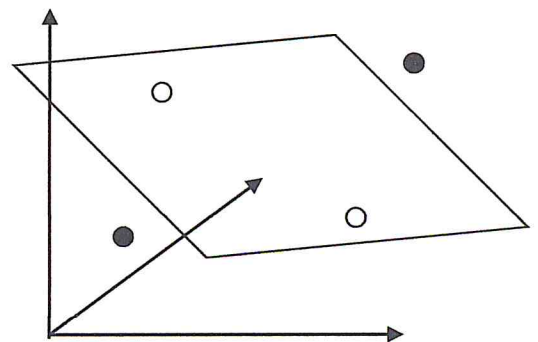
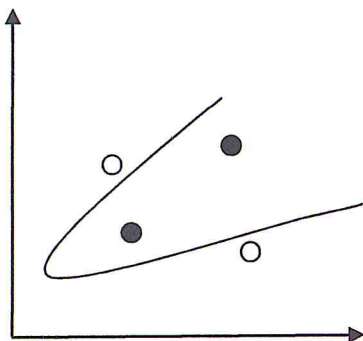


Figure 2.9 : Rôle de la transformation de l'espace par la fonction φ dans la séparation des données.

Alors le problème SVM précédemment décrit, utilise le produit scalaire des données projetées dans R^3 par φ (image des données par φ dans un espace de dimension trois dans cet exemple), et donnerait la fonction linéaire dans R^3 correspondante à une fonction dans R^2 .

Cette transformation (projection), à un espace de dimension trois, est acceptable pour une dimension de données d'entrée inférieure ou égale à deux. Donc pour un cas polynomial, le problème peut devenir plus complexe en temps de calcul. Pour cela il n'est pas nécessaire

d'expliciter la transformation φ à condition d'exprimer le produit scalaire $\langle x_1, x_2 \rangle$ dans un espace H à l'aide d'une fonction k appelée noyau [26] :

$$k(x_1; x_2) = \langle \varphi(x_1), \varphi(x_2) \rangle \quad (2.22)$$

$f(x)$ devient :

$$f(x) = \sum_{i,j=1}^m \alpha_i^* y_i \langle x_i, x_j \rangle + b^* \quad (2.23)$$

Évolué :

$$f(x) = \sum_{i,j=1}^m \alpha_i^* y_i k(x_i, x_j) + b^* \quad (2.24)$$

b) Exemples de noyaux

Nous donnons les noyaux les plus utilisés [25]:

Polynomiale

$K(x, y) = (x \cdot y + c)^d$ La puissance (d) est déterminée *a priori* par l'utilisateur.

Fonction à base radiale

$\exp\left(-\frac{\|x-y\|^2}{2(\delta)^2}\right)$ L'écart type, $(\delta)^2$ commun à tous les noyaux, est spécifié *a priori* par l'utilisateur.

➤ Apprentissage

- ✓ Cette étape a pour objectif de construire un modèle d'apprentissage avec SVMtrain, à partir d'un ensemble d'attributs extraits dans l'étape précédente. Cela revient à déterminer les paramètres d'une fonction **RBF** qui sont : Nu et Gamma dans le but de minimiser l'erreur commise sur l'ensemble des données d'apprentissage. Il s'agit de trouver le plan de séparation optimale (marge maximale) (Consulté l'annexe n°1 pour plus de détails sur ce formalisme).
- ✓ Ce modèle d'apprentissage est utilisé par la suite pour prédire des observations non connues.
- ✓

➤ Teste (détection des CF)

- ✓ Pour prédire l'appartenance d'un nouvel exemple de donnée acoustique (classe positive : CF, classe négative : Non CF), nous avons utilisé SVMpredict. L'étiquette +1 est attribuée à la classe positive et -1 à la classe négative.

2.4.2.2 Positionnement du problème

La sélection des paramètres du modèle SVM est une étape importante. Les paramètres de SVM qui doivent être ajustés pour une efficacité de détection sont :

- ✓ Gamma : paramètre en entrée pour le modèle d'apprentissage dans la fonction du noyau (kernel).
- ✓ Nu : paramètre en entrée pour le modèle d'apprentissage pour traiter un problème à deux classes. Classe positive et négative.

Une fois l'optimisation de ces paramètres est effectuée, la méthode doit théoriquement fonctionner dans des situations réelles.

2.6 Conclusion

Dans ce chapitre, nous avons exposé les processus de détection et reconnaissance des événements impulsifs. Nous avons présenté aussi un état de l'art sur le problème de reconnaissance des CF. Nous trouvons que parmi les méthodes d'apprentissage existant dans la littérature, les modèles SVMs sont les plus adéquats pour notre problème. Par la suite, nous nous sommes focalisés sur la présentation de deux systèmes différents pour la reconnaissance des CF. Dans le chapitre suivant, nous présentons les résultats expérimentaux ainsi que leurs interprétations.

Chapitre 3 : Expérimentations, résultats et interprétations

3.1 Introduction

Dans ce chapitre, nous présentons les résultats expérimentaux obtenus en utilisant les systèmes de reconnaissance proposés dans le chapitre précédent. Nous évaluerons ces systèmes en utilisant les métriques suivantes : Taux de vrais positifs (TVP), Taux de vrais Négatifs (TVN), le Taux de Faux Positifs (TFP), le Taux de Faux Négatifs (TFN), le taux de bonne classification (TBC), et la courbe ROC.

Nous présentons dans la première partie les logiciels de développement utilisés. Dans la deuxième partie, nous présentons les résultats expérimentaux. Par la suite, nous exposons l'évaluation des systèmes. Nous terminons ce chapitre par une conclusion.

3.2 Logiciels de développement

3.2.1 Environnement MATLAB

Nous avons utilisé dans notre projet la version 2009 (Matlab R2009a), pour le prototypage et le test des solutions proposées (Figure 3.1).

Le logiciel MATLAB est un environnement informatique numérique et un langage de programmation. Un grand nombre de fonctions dédiées au traitement de signal et aux statistiques est disponible sous cette version. Le choix de l'environnement MATLAB a été fait en raison de sa rapidité. Il permet d'effectuer des opérations mathématiques, de manipuler des matrices, de faire des graphiques.

Les Toolboxes qui ont été utilisées sont :

- ❖ **Libsvm** [17] « Library for SVM » : c'est une bibliographie externe pour les SVMs. Nous l'avons téléchargé et installé sous MATLAB.
- ❖ **Wavelet Toolbox** [21]: C'est une bibliographie pour analyser et synthétiser des signaux, des images et des données qui présentent un comportement régulier ponctué de changements abrupts. Cette bibliothèque est intégrée dans l'environnement MATLAB 2009.

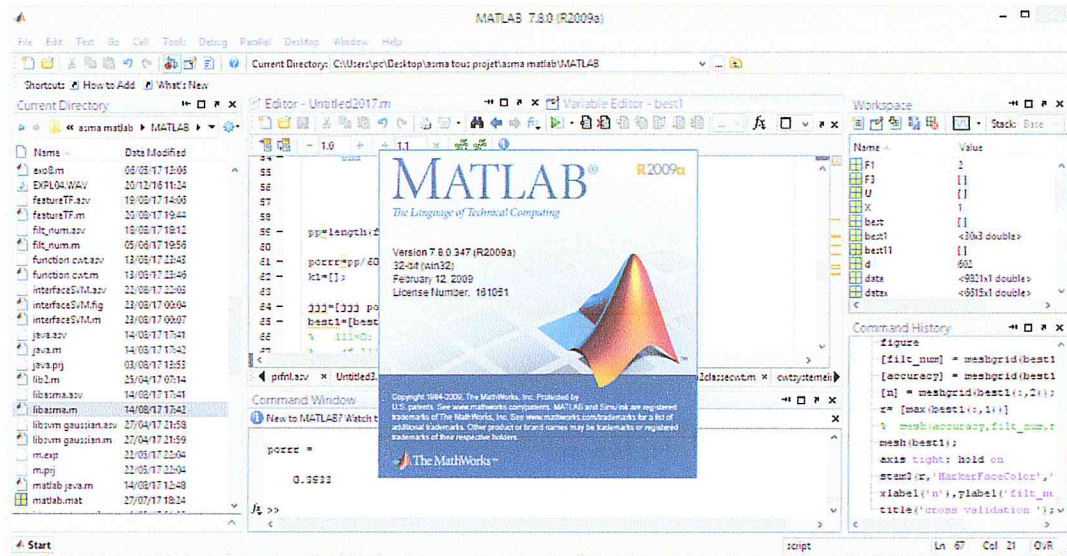


Figure 0.1 : MATLAB R2009a

3.2.2 GoldWave

GoldWave un logiciel sous Windows qui permet la lecture et l'enregistrement de fichiers audio sous formats divers (.WAV, .MP3, .Wma, .Aiff, etc...). Il est doté de plusieurs effets acoustiques comme Doppler, Dynamics, Echo, Filter, et autres. Il permet aussi d'éditer, copier, coller et convertir les fichiers audio [22]. Dans notre projet, nous avons utilisé GoldWave pour le prétraitement et la construction du corpus acoustique. Il a été utilisé aussi pour éliminer les zones de silence.

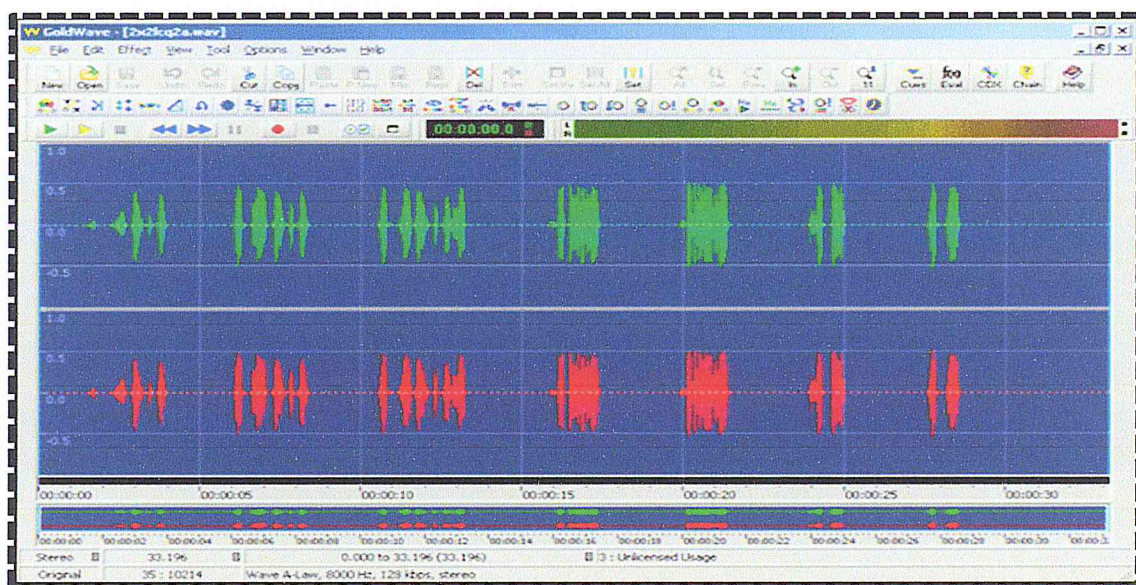


Figure 0.2 : GoldWave

3.3 Métriques d'évaluation

3.3.1 Matrice de confusion

Nous allons évoquer, dans cette partie, la matrice de confusion. C'est un tableau à deux dimensions (Tableau 3.1) qui rassemble les erreurs de prédiction pour un problème à deux classes [18].

Tableau 3.1 : Matrice de confusion [18].

		Conditions réelles	
		Condition positive	Condition négative
Décisions (Prédiction)	Décision positive	Vrai Positifs (VP)	Faux positifs (FP)
	Décision négative	Faux Négatifs (FN)	Vrai Négatifs (VN)

Condition positive : le son est réellement un CF,

Condition négative : le son est réellement non CF,

Décision positive : le test est positif,

Décision négative : le test est négatif,

Vrai Positifs (VP) : les CF capturés au test sont classés des CF,

Faux positifs (FP) : Les autres sons impulsifs (Non CF) capturés au test sont classés comme CF,

Faux Négatifs (FN) : les CF capturés au test sont classés non CF,

Vrai Négatifs (VN) : Les autres sons impulsifs (Non CF) capturés au test sont classés comme non CF,

Taux de vrais positifs (TVP) :

$$TVP = \frac{VP}{VP+FN} \quad (3.1)$$

Taux de vrais Négatifs (TVN) :

$$TVN = \frac{VN}{VN+FP} \quad (3.2)$$

Taux de Faux positifs (FP)

$$TFP = \frac{FP}{FP+VN} \quad (3.3)$$

Taux de Faux Négatifs (FN) :

$$TFN = \frac{FN}{FN+VP} \quad (3.5)$$

Taux de bonne classification ou l'exactitude (TBC) :

$$TBC = \frac{VP+VN}{VP+FN+VN+FP} \quad (3.6)$$

3.3.2 Courbe ROC

La courbe ROC est un moyen de comparaison des tests de diagnostic. Dans une courbe caractéristique de fonctionnement du récepteur (ROC), le taux vrai positifs (Sensibilité) est tracé en fonction du taux de faux positif (spécificités), pour différents points de coupure. Chaque point de la courbe ROC représente une paire de sensibilité / spécificité correspondant à un seuil de décision particulier. Par conséquent, plus la courbe ROC est proche du coin supérieur gauche, plus la précision globale du test est élevée [19].

3.3.3 Validation croisée

Parmi les méthodes d'évaluation et d'estimation de vraisemblance du modèle, la validation croisée est la meilleure. Nous avons choisis la technique (k-fold cross-validation/ k=10) [20], (Figure 3.3). Les étapes de cette technique sont :

- décomposer les données en k sous-ensembles de taille n/k . n est la taille de l'ensemble de données.
- Faire l'apprentissage sur k-1 ensembles (folder) de données et tester sur 1 seul folder.
- Répétez k fois les étapes précédentes et prenez la moyenne des TBCs.

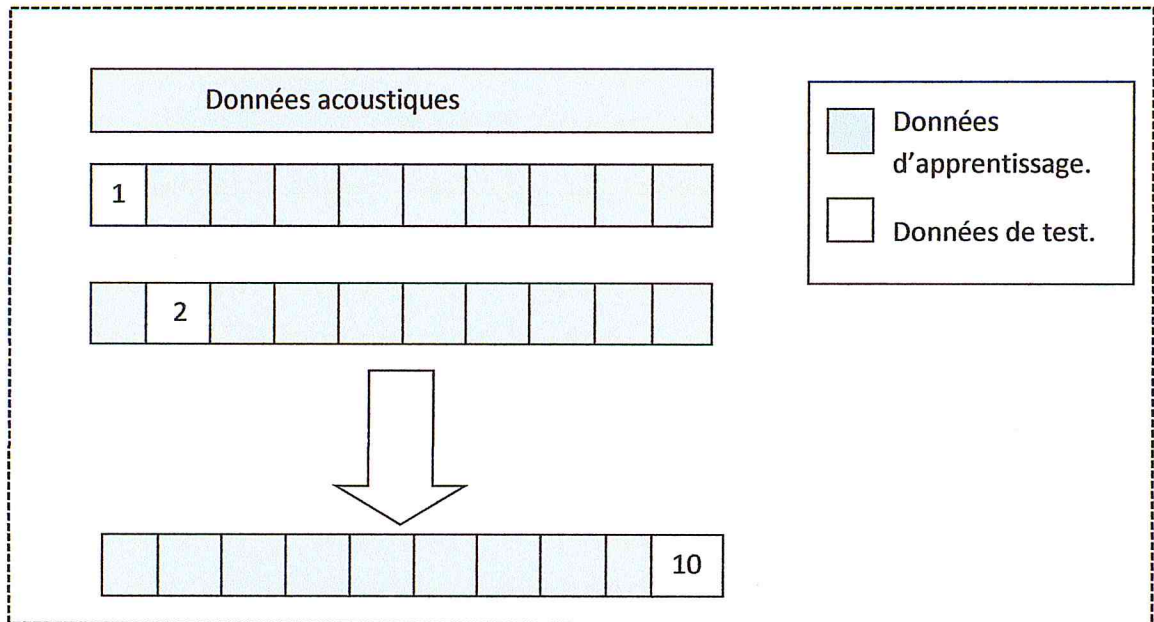


Figure 3.3 : Principe de la validation croisée.

3.4 Corpus de tests

Nous avons téléchargé nos sons à partir du site Sounddogs [23]. C'est une base de données audio contenant environ 694 395 fichiers englobant toutes les catégories des sons (sons des guitares, la nature, fermeture des portes, explosions, armes à feu, les animaux, les cris humains et autres).

Tous les signaux audios sont échantillonnés à 11025 Hz avec une quantification de 16 bits.

Nous avons téléchargé des sons de CF (pistolets, revolvers et carabines) : la classe positive.

Nous avons téléchargé aussi des sons de claquements de portes, des aboiements de chiens et des cris humains : classe négative.

Au total, notre base de données représentait 1200sons audio comprenant 600 CF et 600 évènements dangereux.

Nous avons utilisé une durée de 0.6 seconde pour chaque signal de notre base de données, dans le but d'avoir une matrice de coefficients (TOC) avec une taille fixe est de durée suffisante pour la reconnaissance des CF.

Nous avons utilisé pour la reconnaissance des CF 600 SI.

300 CF :

- ✓ Pistolets
- ✓ Révolvers
- ✓ Carabines

300 autres sons impulsifs :

- ✓ Claquements de portes,
- ✓ Aboiement de chiens,
- ✓ Crie humains.

3.5 Étude expérimentale

3.5.1 Premier système

- ✓ Les paramètres de ce système qui doivent être ajustés pour une efficacité de reconnaissance sont :
 - ✓ La fréquence basse « $F_{\text{Mel_basse}}$ »,
 - ✓ La haute fréquence « $F_{\text{Mel_haute}}$ »,
 - ✓ Le nombre de trame « N »,
 - ✓ Le nombre d'échelles « M ».
- $F_{\text{Mel_basse}}$: Vu que la gamme de la perception humaine est comprise entre 20Hz et 20kHz. Nous avons choisi la valeur 20 Hz comme fréquence basse.
- $F_{\text{Mel_haute}}$: Elle est fixée à $(\frac{11025}{2})$ Hz vu que la F_e de notre corpus est égale à 11025 Hz.
- N : C'est le nombre de colonne de la matrice de l'énergie. Nous avons utilisé la plage de variation suivante lors de l'optimisation par Grid search : $N= 2$, $N= 3$, $N= 4$ et $N= 5$, pour satisfaire les contraintes suivantes : La valeur de ce paramètre doit être assez faible pour une exécution rapide (Real Time), avec une complexité réduite. Donc il faut au moins sélectionner deux trames en vue de prendre en compte l'évolution temporelle d'une analyse temps- fréquence (temps échelle).
- M : c'est le nombre d'échelles. Les valeurs que nous avons utilisées sont : $M=5$, $M=6$, $M=7$, $M=8$, $M=9$, et $M=10$). Ce choix est motivé pour les mêmes raisons citées ci-dessus.

3.5.2 Le deuxième système

Nous avons présenté les paramètres de l'étage d'extraction d'attributs dans le chapitre précédent, nous avons extrait un vecteur V_{eng} de taille $M \times N$, où ' N ' est le nombre de trames et ' M ' est le nombre d'échelles. L'ensemble des données d'apprentissage sont représentés par

une matrice d'apprentissage construite à partir des vecteurs d'énergies Veng. Nous avons fourni cette matrice comme entrée pour le classifieur SVM.

Pour la sélection du modèle (Model selection), nous avons optimisé les paramètres Gamma et Nu à base de Grid search.

- **Gamma** : nous avons utilisé ces valeurs pour évaluer ce paramètre (Gamma=0.1, Gamma=0.2, Gamma=0.3, Gamma=0.4, Gamma=0.5, Gamma=0.6, Gamma=0.7, Gamma=0.8, Gamma=0.9), la valeur par défaut de ce paramètre est (Gamma_default=1/nombre d'attributs). Donc notre choix s'est fait par rapport à cette valeur.
- **Nu** : Nous avons choisi d'utiliser Nu comme paramètre à optimiser et non pas le paramètre C vue que ce dernier vari de 0 jusqu'à l'infinie. Par ailleurs, le Nu est normalisé entre 0 et 1. La valeur par défaut de Nu est (0.5). Les valeurs possibles de Nu sont : (Nu=0.1, Nu=0.2, Nu=0.3, Nu=0.4, Nu =0.5).

3.6 Application développée

Nous avons implémenté nos systèmes sous Matlab R2009a (versions off line).

3.6.1 Premier système

Ce système est représenté par le GUI suivant (Figure 3.4).

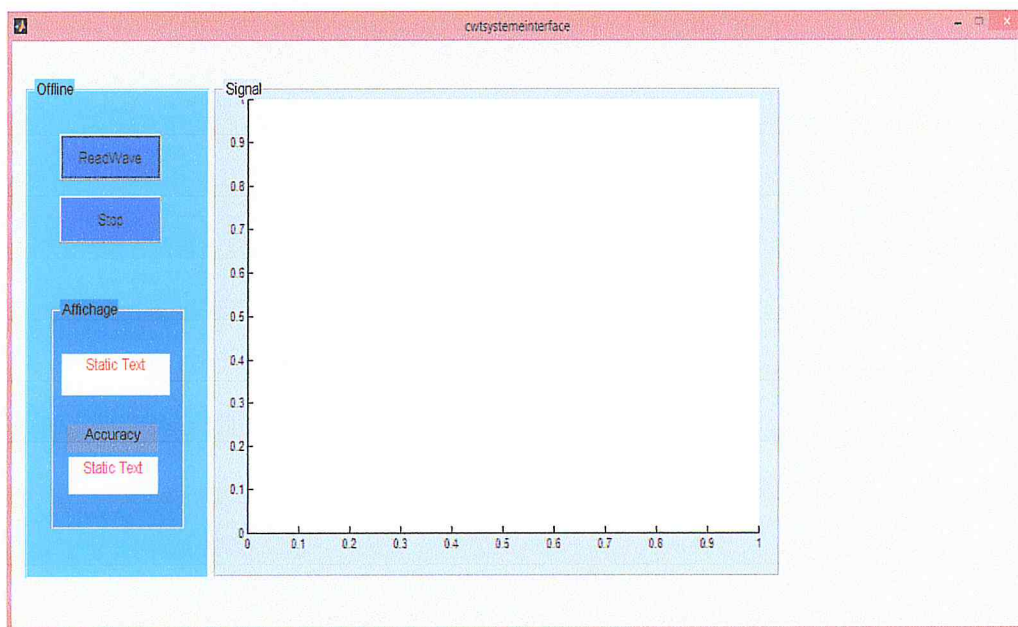


Figure 3.4: Premier application

3.6.2 Deuxième système

Ce système est représenté par (Figure 3.5).

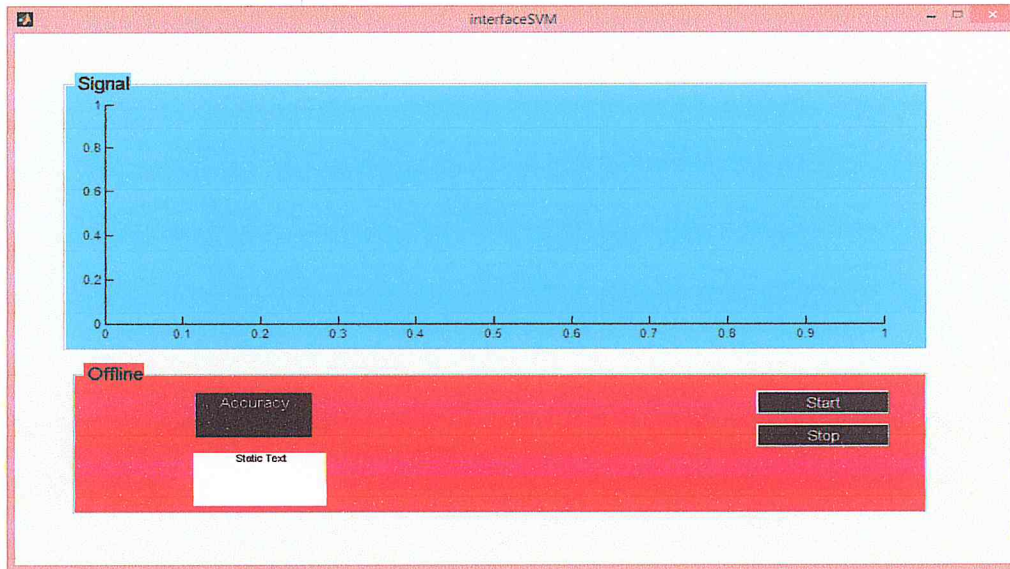


Figure 3.5 : Deuxième application

3.7 Évaluation des systèmes

3.7.1 Première système

Pour l'évaluation de performance du premier système sur le corpus élaboré, nous utilisons la matrice de confusion, le TVP, le TVN et le TBC.

Une approche de Grid search est utilisée pour trouver les meilleures valeurs des paramètres de ce système.

D'après les résultats, nous avons obtenu les valeurs suivantes :

$N = 2$, et $M = 5$. Ce choix se fait par rapport aux meilleurs TBCs (0.77 %)

Pour valider notre choix, nous avons utilisé un autre ensemble de sons qui n'a pas été utilisé dans notre première expérience. Nous avons trouvé que le TBC est égal à 0.761% (Le choix de M et N est correct).

best1 <24x3 double>								
	1	2	3	4	5	6	7	8
1	0.8900	2	5					
2	0.8733	3	5					
3	0.8467	4	5					
4	0.8167	5	5					
5	0.8267	2	6					
6	0.7667	3	6					
7	0.7367	4	6					
8	0.6900	5	6					
9	0.7500	2	7					
10	0.6900	3	7					
11	0.6400	4	7					
12	0.5900	5	7					
13	0.6667	2	8					
14	0.6267	3	8					
15	0.5900	4	8					
16	0.5167	5	8					
17	0.6300	2	9					
18	0.5767	3	9					
19	0.5467	4	9					
20	0.4633	5	9					
21	0.5867	2	10					
22	0.5400	3	10					
23	0.4867	4	10					
24	0.4333	5	10					
25								
26								

Figure 3.6 : Étude paramétrique

- **Matrice de confusion**

Nous avons calculé la matrice de confusion (sous forme de taux d'erreurs en %)

Tableau 3.2 : Matrice de confusion de système (1)

		Conditions réelles	
		CF (Positive)	Non CF (Négative)
Décisions	Décision de CF (positive)	0.873 % TVP	0.33 % TFP
	Décision de Non CF (Négative)	0.126 % TFN	0.67 % TVN

Nous avons obtenu un TBC=0.77, d'après les résultats de matrice de confusion.

3.7.2 Deuxième système

Pour l'évaluation des performances du deuxième système en utilisant le corpus élaboré, nous nous sommes basés sur la validation croisée. Nous avons utilisé les métriques suivantes : la matrice de confusion, le TVP, le TVN, le TBC et la courbe ROC.

- **Cross-validation (validation)**

Nous avons utilisé 10-fold cross validation comme technique de sélection et grid search comme approche d'optimisation. Nous avons trouvé les valeurs optimales suivantes :

$\text{Nu}=0.2$, $\text{Gamma}= 0.45$ pour un TBC = 92,49%. Ce choix se fait par rapport à la moyenne des 10 résultats que nous avons trouvés.

- **Courbe ROC**

La courbe ROC pour ce système est représenté sur la figure 4.6

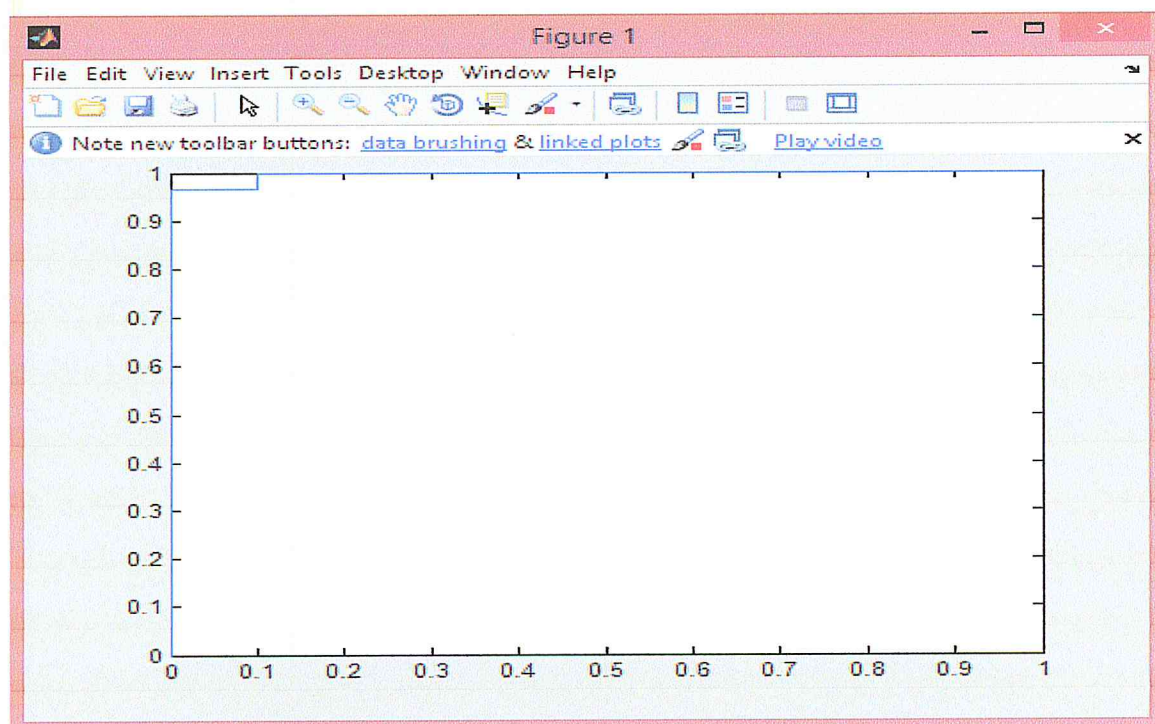


Figure 3.7 : Courbe ROC

- **Matrice de confusion**

Tableau 3.3 : Matrice de confusion du système (2)

		Condition réel	
		CF	Non CF
Décision	Décision de CF	1% TVP	0.033% TFP
	Décision de Non CF	0% TFN	0.967% TVN

Nous avons obtenu une TBC= 0.9833, d'après les résultats de matrice de confusion.

3.7 Conclusion

Nous avons étudié deux systèmes différents dans ce chapitre. Dans la première partie, nous avons présenté les logiciels de développement. La deuxième partie de ce chapitre concerne l'étude expérimentale pour l'optimisation des paramètres en vue d'assurer une bonne performance de détection des CF et cela pour les deux systèmes proposés. Dans la partie d'évaluation, nous avons trouvé les paramètres suivants : (i) Pour le premier système les paramètres sont $N=2$, et $M=5$. (ii) Pour le deuxième système nous trouvons que $N=2$ et $M=5$, $\gamma=0.45$, $\nu=0.2$. Les résultats trouvés sont : (i) pour le premier système $TBC=0.77$, $TVP=0.873$, $TVN=0.67$, $TFP=0.33$, $TFN=0.126$ (ii) et pour le deuxième système, nous avons comme résultats : $TBC=0.983$, $TVP=1$, $TVN=0.967$, $TFP=0.033$, $TFN=0$.

Chapitre 4

Conclusions et travaux futures

4.1 Conclusions

L'objectif initial de ce travail est de concevoir un système de détection des coups de feu en vue de la mise en œuvre d'un système de surveillance audio. Précisément, il s'agit de différencier entre les sons des coups de feu et les autres sons impulsifs de l'environnement.

Dans notre travail, nous nous sommes basés sur la modalité sensorielle afin de détecter les coups de feu. Nous avons proposé deux systèmes différents pour la reconnaissance des coups de feu : (i) Le premier système est basé sur un seul étage de traitement, (ii) le deuxième système est composé de deux étages de traitement, l'extraction d'attributs et la classification.

L'étage de traitement du premier système consiste en l'utilisation de la transformée en ondelettes continue (TOC) en vue de la réalisation des règles de décision.

Pour le deuxième système, l'extraction d'attributs est basée aussi sur la TOC pour la génération de plusieurs mesures énergétiques. La classification est basée sur les Support Vector Machines (SVMs).

La construction du corpus sonore est un facteur important pour l'élaboration d'un système de reconnaissance des coups de feu (facteur de satisfaction de la solution). Nous avons utilisé trois types d'armes à feu : revolvers, pistolets et carabines. Pour chaque type, nous avons exploité 100 sons. Notre corpus contient une totalité de 600 sons impulsifs : 300 coups de feu et 300 autres événements impulsifs : cris humains, fermeture des portes et aboiement de chiens.

Afin d'obtenir une procédure de détection des coups de feu ayant de bonnes performances de reconnaissance ainsi qu'une complexité réduite, nous avons optimisé nos paramètres algorithmiques avec une approche de Grid Search.

Pour le premier système, nous avons trouvé que le nombre de trame N doit être égal à 2. Cependant, le nombre d'échelles M doit être à 5. Pour le deuxième système nous avons trouvé : (i) les paramètres de l'étage d'extraction d'attributs $N=2$ et $M=5$, et (ii) les paramètres de l'étage de classification (model selection) : $\gamma=0.45$, $\nu=0.2$.

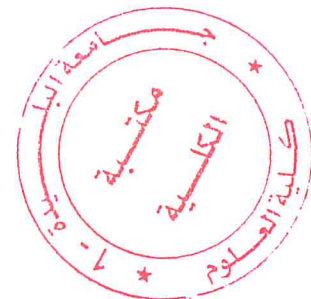
D'après l'évaluation des performances, les résultats trouvés sont : (i) pour le premier système : $TBC=0.77$, $TVP=0.873$, $TVN=0.67$, $TFP=0.33$, $TFN=0.126$ (ii) et pour le deuxième système, nous avons comme résultats : $TBC=0.983$, $TVP=1$, $TVN=0.967$, $TFP=0.033$, $TFN=0$.

D'après les résultats trouvés, nous concluons que :

- (i) Le premier système a une complexité réduite par rapport au deuxième système, vu qu'il se base sur des règles de décision. Ce système est plus approprié à la reconnaissance en temps réel des CF. De plus, il n'utilise pas un étage de classification à base des machines d'apprentissage,
- (ii) Le deuxième système est plus complexe mais possède de meilleures performances,

4.2 Travaux futures

- Implémentation d'un système de détection de CF en temps réel en utilisant la solution décrite dans ce travail,
- Logiciel sous JAVA,



ANNEXES

Annexe n° 01 : Classification des types d'armes

Il est facile de se méprendre sur les classifications d'armes. Plusieurs droits nationaux et internationaux de contrôle utilisent des termes différents pour définir et classer les types d'armes.

Pour cela, puisqu'une certaine connaissance de base des armes à feu est nécessaire, nous avons conduit une étude sur les types d'armes. Selon notre analyse, nous pouvons dire qu'il existe deux catégories générales d'armes légères : les armes de poing et les armes d'épaule.

1. Armes de poing

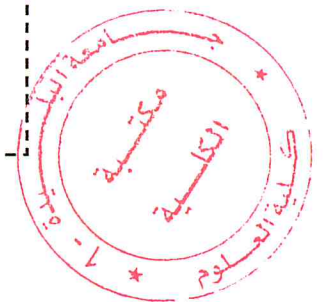
C'est une arme à feu utilisable à une main. Elle présente un encombrement et une masse réduite. Elle peut être portée en permanence. Il existe deux types basiques d'armes de poing : les revolvers et les pistolets [3].

A. Revolver

Un revolver est une arme avec un cylindre/barillet tournant qui contient un certain nombre de chambres. Chaque chambre englobe une cartouche qui peut être successivement alignée avec le canon et la tige de tir, et déchargée à travers un seul barillet [4].



Figure 1 : Révolver



B. Pistolet

Le pistolet (précisément pistolet semi-automatique) est une arme de poing dont la chambre est intégrée au canon ou alignée avec lui en permanence. Les pistolets ont des taux de tir plus élevés et peuvent être chargés plus rapidement que les revolvers [5].

2. Armes d'épaule

Une arme d'épaule est une arme à feu tenue à l'épaule. Elle est plus précise et plus puissante qu'une arme de poing. Cependant, une arme d'épaule est plus encombrante et plus lourde [6]. On distingue deux catégories principales : (i) les armes à canon lisse et (ii) les armes à canon rayé.

A. Armes à canon lisse

Les armes à canon lisse selon la vue de la réglementation française et du droit, sont les armes dont l'alésage du canon est de section circulaire et ne peut donner aucun mouvement de rotation à un projectile unique ou multiple. On peut distinguer deux sous catégories : (i) les fusils de chasse et (ii) les fusils à pompe [6].

A.1 Fusils de chasse (Shotgun)

Un fusil de chasse est conçu pour tirer plusieurs granulés du barillet. On peut faire la distinction entre deux sous-classes de fusil : (i) fusil à seul canon, et (ii) fusil à double canon. Les fusils de chasse à double canon peuvent être positionnés les uns sur les autres (sur et sous) ou superposés (côte à côte) [4] [7].



Figure 2: Fusil de chasse.

A.2 Fusils à pompe

Un fusil à pompe est un fusil dans lequel le chargement se fait par un mouvement d'une poignée (située sous le canon) d'avant en arrière. Les cartouches sont creusées dans un chargeur tabulaire situé sous le canon [6].



Figure 3 : Fusil à pompe

B. Armes à canon rayé

Au regard de la réglementation française et du droit, les armes à canon rayé sont des armes où l'alésage du canon présente une ou plusieurs rayures ce qui confère une rotation à un projectile unique ou multiple. La rotation est créée pour stabiliser le projectile [3].

Il existe trois sous catégories d'armes à canon rayé : (i) les carabines à répétition manuelle (Rifle), (ii) les mitrailleurs (Submachine Guns) et (iii) les pistolets mitrailleurs (Machine Pistols) [3].

B.1 Carabines à répétition manuelle (Rifle)

Les carabines sont des fusils d'épaule munies d'une crosse. Elles utilisent les mêmes munitions que celles des fusils. Néanmoins, la précision de leur tir est moins exacte que les fusils ordinaires [5]. Généralement, les carabines possèdent une seule chambre de tir intégrale [8].



Figure 4 : Carabine à répétition manuelle (Rifle)

B.2 Mitraillettes (Submachine Guns)

Une mitrailleuse est une arme conçue pour être tirée de l'épaule. Elle peut être configurée en mode de tir coup par coup ou en rafale (automatiquement ou sélectif). La plupart des mitrailleuses sont alimentées par des ceintures ou des magazines. Les mitrailleuses sont utilisées par des unités spéciales des forces armées et de sécurité [4] [7].

B.3 Pistolet-mitrailleur (Machine Pistols)

Les pistolets mitrailleurs (pistolet-automatique) sont des armes conçues pour tirer des rafales qui ne dépassent pas les trois coups. Ils offrent la possibilité de choisir deux modes de tir : (i) automatique ou (ii) semi-automatique. Les pistolets mitrailleurs sont utilisés principalement par des unités spéciales de la police ou des gardes de sécurité [5].



Figure 5: MP5 (un pistolet-mitrailleur).

Tableau 1 : Vitesses du museau de plusieurs armes (29).

Arme	Calibre	Vitesse du museau (m / s)
M16 fusil	5.56 mm	1000
Remington 30-06 fusil	30.06	960
AK-47 Sous-machine pistolet	5.45 mm	700
Revolver Magnum	0.500	550
Revolver Magnum	0.44	390–490
Beretta Modèle 96D	40	445
Pistolet	9 mm	280–350
Heckler& Pistolet tactique Koch USP	0.45	260
Revolver	0.22	200

Bibliographie

- [1] M. Cristani, A. Trucco et M. Vittorio, «Audio Surveillance:A systématique réview.,» vol. 48, no.4,30 Septembre 2014.
- [2] S. Lecomte, "Classification Partiellement Supervisée par SVM Application à la Détection d'Evènements en Surveillance Audio," Thèse de doctorat, Université de Technologie de Troyes, France, 2013.
- [3] Professionnels du droit et de la finance, la direction d'Eric Roig, diplômé d'HEC, «Définition juridique,» juin 2014. [En ligne]. Available: <http://droit-finances.commentcamarche.net/faq/26695-arme-a-feu-definition-juridique>.
- [4] J. Brian, *Handbook of firearms and ballistics*, J. Wiley, Éd., Hong Kong: Wiley-Blackwell, 2008.
- [5] UNREC, «Identification des armes légères et de petit calibre et de leurs munitions,» vol. 20, no.2,19 06 2009.
- [6] A. Gallussaire, M. Bonfanti et F. shudz, *Expertise des armes à feu et des éléments de munitions dans l'investigation criminel*, Lausanne, Éd., Presses polytechniques et universitaires romandes, 2002.
- [7] UNREC, «Identification des armes légères et de petit calibre et de leurs munitions,» vol. 20, no.1,19 06 2009.
- [8] J. Vincent et M. Dimaio, *Gunshot wounds*, New York: CRC press LLC, 1999.
- [9] J. C. Swallow, J. Hemingway et P. Yung, *Acoustique des champs de tir*, 1999.
- [10] A. L. L. Ramos, S. Holm, S. Gudvangen et R. Otterlei, «A Spectral Subtraction based algorithm for real-time Noise cancellation with application to gunshot acoustics,» *INTL.journal of electrics and telecommunications(jet)*, Vols. 59, n° 11, p. 93–98, 2013.
- [11] A. M. Borzino, J. A. Apolinario, M. L. DeCompos et C. L. Pagliari, *Gunshot signal enhanchement for DOA estimation and weapon recognition*, Rio de Janeiro, 2013.
- [12] A. Talal, U. Momin and A. Muhammad, "Improving Efficiency and Reliability of Gunshot Detection Systems," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'13)*, Vancouver, BC, Canada, 2013.
- [13] A. Rabaoui, H. Kadri, Z. Lachiri and N. Ellouze, "One-Class SVMs Challenges in Audio Detection and Classification Applications," *EURASIP Journal on Advances in Signal Processing*, 2008.
- [14] J. Dumas, *L'Analyse temps-fréquence*, Groupe MVI technologie,, 2001.

- [15] K. Mahkonen, *Mel-frequency cepstral coefficients (MFCCs) and filter banks*, Pasi Pertila, 2015.
- [16] J. Landerie, *Boîte à Outil sous Matlab*, Université de Bourgogne, 2002.
- [17] C.-C. Chang et C.-J. Lin, «LIBSVM,» [En ligne]. Available: <http://www.csie.ntu.edu.tw/~cjlin/libsvm/oldfiles/index2.0.html>. [Accès le 2017].
- [18] Y. Oufella, *Evolution du concept de front roc et combinaison de classificateur*, ,Memoire du master, Université de Rouen UFR des scientifique et thechnique, 2008.
- [19] M. H. Zweig et G. Campbell, «Receiver-Operating Characteristic (ROC) plots: A F undamental Evaluation Tool In Clinical Medcine,» vol. 39, no.4,Avril 1993.
- [20] S. Sylvain Arlot, *La validation croisée*, <https://hal.archives-ouvertes.fr/hal-01485508>, 2017.
- [21] «MathWorks,» [En ligne]. Available: <https://www.mathworks.com/help/wavelet>. [Accès le 2017].
- [22] «GoldWave digital audio Editing Software,» [En ligne]. Available: <http://www.goldwave.com>. [Accès le 2017].
- [23] «Sounddogs,» [En ligne]. Available: <https://retired.sounddogs.com/about-us.asp>. [Accès le 2017].
- [24] G. Lebruni, C. Charrier et O. Lezoray, «Réduction du temps d'apprentissage des SVM par Quantification Vectorielle,» Saint-Lô.
- [25] A. Cornuéjols, «Une nouvelle méthode d'apprentissage :Les SVM. Séparateur à vaste marge,» Université de Paris-Sud,France,juin 2002.
- [26] F. Lauer et G. Bloch, *Méthodes SVM pour l'identification*, Université Nancy,France, 2006.

