

الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire

Ministère de l'Enseignement Supérieur
et de la Recherche Scientifique

Université Akli Mohand Oulhadj - Bouira -

Tasdawit Akli Muḥend Ulḥağ - Tubirett -



وزارة التعليم العالي والبحث العلمي
جامعة أكلي محمد أولحاج
- البويرة -

Faculté des Sciences et des Sciences Appliquées
Département de Génie Electrique

كلية العلوم و العلوم التطبيقية
قسم الهندسة الكهربائية

Option : Technologies des Télécommunications

Thème :

**Tatouage audio utilisant le
masquage perceptuel**

Réalisé Par:

-BENHACICENE Wafa

-MADDEHI Meriem

Date de soutenance : 26/09/2017

Devant le jury composé de :

Mr.MOUDACHE Saïd

Mr.SAIDI Mohammed

Mr.SAOUD Bilal

Mr.OUILI Kahina

Président

Rapporteur

Examineur

Examineur

Année Universitaire: 2016-2017

Dédicaces

Je dédie Ce modeste travail à:

A mes parents. Aucun hommage ne pourrait être à la hauteur de l'amour Dont ils ne cessent de me combler. Que dieu leur procure bonne santé et longue vie.

*A celui que j'aime beaucoup et qui ma soutenue tout au long de ce projet : mon cher **ADEL** et bien sur A mes frères « **ABDELHAK, NESREDDINE et ABDELFATTAH** », et mes sœurs « **KARIMA, FAYZA, SAMIRA, MAHDIA et NAIMA** », et la belle sœur **NASSIRA**, sans oublié ma nièces **MARIA** et mes neveux **IYAD** et **SID AHMED***

*A tout ma famille ET mes amis « **NASSIRA, AMINA, ZAHRA, SOUHILA et BAYA** ».*

*A mon binôme **MERIEM**.et toute la famille **MADDAHI**.*

Et à tous ceux qui ont contribué de prés ou de loin pour que ce projet soit possible, je vous dis Merci.

Wafa

Dédicace :

Je dédie ce travail à ma chère mère pour tout l'amour et la patience qui ma offert, pour tous les encouragements et le soutien qui m'a bercé au long de ma vie, je tiens à dire que ce diplôme est le couronnement de ton effort.

A mon cher père Allah yarhmo

*A mon très cher mari **Razki***

*A mes sœurs **Aicha, Fati, Hada, Khadidja***

*A mes frères **Lounes, Ahmed et Abdo***

*A tous mes ami(e) s et mes collègues (**Biba, Zahra, Wafa, Rafika, Aicha, Wawa, Soso, Hanane, Nana,**).*

*A toute ma famille «**MEDDAHI et DABAGHI**»*

MERIEM

Remerciement

Après avoir remercié Dieu,

J'exprime ma profonde reconnaissance à monsieur SAIDI Mohamed enseignant à l'université de Bouira pour avoir proposé et dirigé ce travail efficacement.

Mes remerciements s'adressent en premier lieu aux membres du jury pour l'intérêt qu'ils ont porté à mon travail.

Ma gratitude s'adresse également à l'équipe des enseignants de notre première promotion Master2 de Technologie de télécommunication à université de Bouira, pour tout le savoir qu'ils nous ont transmis.

Merci également à tous mes collègues de télécommunication pour leur gentillesse, compréhension et ambiance.

Enfin un grand merci à tous mes amis qui m'ont toujours aidé.

Résumé

L'avènement des documents sous format numérique offre des possibilités innovantes dans le domaine d'embarcation des données ; un document hôte, dans notre cas un signal audio, peut cacher une information numérique. La technique d'embarcation ou d'enfouissement est alors appelée tatouage.

Notre travail consiste à mettre en place une procédure de tatouage permettant d'intégrer une protection de droit d'auteur pour des données audio numériques, par modification directe des échantillons audio. Cette procédure exploite directement le masquage perceptuel, temporel et fréquentiel, pour garantir que le filigrane numérique (tatouage) soit inaudible et robuste.

Les Mots clés : Tatouage audio, Masquage perceptuel, psycho-acoustique

Liste des tableaux

<i>Tableau II.1</i> : Bandes critiques.....	17
Tableau IV.1 : Résultats des tests de temps de calcul sure différents signaux audio échantillonnés à 32KHz.....	44

Table des figures

Figure 1.1 : Techniques de la sécurité de l'information.....	3
Figure 1.2 : Schéma général d'un système de dissimulation d'information.....	4
Figure 1.3 : Système de tatouage de référence en présence de perturbations externes.....	7
Figure 1.4 : Schéma Général d'un système de tatouage numérique.....	9
Figure 2.1 : seuil audition absolu.....	16
Figure 2.2 : Masquage fréquentielle	20
Figure 2.3 : Masquage temporel.....	23
Figure 3.1 : Schéma génial du tatouage.....	28
Figure 3.2 : masque fréquentielle.....	29
Figure 3.3 : spectre de puissance du signal.....	30
Figure 3.4 : les composants tonals.....	31
Figure 3.5 : élimination des composants masqués.....	32
Figure 3.6 : seuil audition fréquentielle.....	32
Figure 3.7 : masque temporel.....	33
Figure 3.8 : Principe de signature para pore au signal audio.....	34
Figure 3.9 : bruit pondéré par le masque fréquentiel.....	35
Figure 3.10: masque fréquentiel.....	36
Figure 3.11 : tatouage global du signal audio.....	38
Figure 3.12 : spectre de bruit.....	39
Figure 4.1 : partie voisée du signal tatoué.....	40
Figure4 .2 : partie bruité du signal tatoué.....	41
Figure4.3 : Partie silencieuse du signal tatoué.....	42

Liste des abréviations

BBAG	Bruit Blanc Additif Gaussien
BF	Basse fréquence
CD	Compact Disc
DCT	Discrete Cosinus Transform
DSP	Densité Spectrale de Puissance
FFT	fast-Fourier transform (transformée de Fourier rapide)
FM	Modulation en fréquence
G	Coefficients du filtre $G(f)$ sous forme vectorielle
HF	haute fréquence
IFFT	inverse fast-Fourier transform (transformée de Fourier inverse rapide)
IES	Interférence Entre Symboles
IVDS	Interactive Video and Data Services
ISO	International Standards Organisation
MPA	Modèle Psycho Acoustique
MPEG	Motion Picture Expert Group
Mp3	MPEG 1 Audio Layer 3
N	Taille des fenêtres d'analyse
RI	Réponse impulsionnelle
RMS	Rapport de puissance entre le seuil de Masquage et le Signal audio
SPL	Sound Pressure Level
TEB	Taux d'Erreur Binaire

Unités

ms	méli second
dB	décibels
KHz	Kilo Hertz
Hz	Hertz

Nomenclature

$x(t)$	Signal audio original
$y(t)$	Signal audio tatoué
$\hat{y}(t)$	Signal audio tatoué détérioré
$v(t)$	Signal modulé
$\hat{v}(t)$	Signal audio tatoué détérioré puis blanchi
$W(t)$	Signal de tatouage
F_0	Fréquence porteuse du signal $v(t)$
$H(f)$	Filtre de mise en forme spectrale associé au signal audio original x
$G(f)$	Filtre blanchissant associé au signal audio original x
D	Dictionnaire de modulation d'émission
\hat{D}	Dictionnaire de démodulation de réception

Table des matières

Sommaire

Remerciement.....	I
Résumé	II
Sommaire.....	III
Liste des tableaux.....	IV
Liste des figure.....	V
Liste des abréviations	VI
Introduction générale.....	1

Chapitre I : Généralités sur le tatouage audio

I.1 Introduction.....	2
I.2 Dissimulation de l'information, Stéganographie et Tatouage.....	3
I. 2.1 La Dissimulation de l'information.....	3
I. 2.1.1 Caractéristiques d'un système de dissimulation de l'information.....	4
I. 2.2 La Stéganographie.....	6
I. 2.3 Tatouage ou « Watermarking ».....	7
I.3 Tatouage Numérique.....	8
I. 3.1 Définition de tatouage numérique.....	9
I. 3.2 Caractéristiques du tatouage numérique.....	9
I.4 Le tatouage audio.....	10
I. 4.1 Applications du tatouage audio.....	10
I. 4.2 Objectifs et contraintes.....	11
I.5 Conclusion.....	12

Chapitre II: Modèle psycho-acoustique et masquage perceptuel

II.1 Introduction	13
II.2 Notion de Psycho acoustique	14
II. 2.1 Définition	14
II. 2.2 Seuil d'audition absolu	15

II. 2.3 Bandes critiques	16
II.2.4 La compression audio et la psycho acoustique	18
II.3 Masquage perceptuel	19
II. 3.1 Présentation	19
II.3.2 Masquage fréquentiel	20
II.3.2.1 Seuil de masquage	20
II.3.2.2 NormeISO/MPEG.....	21
II.3.3 Masquage temporel.....	22
II.3.3.1 Transformé d’Hilber.....	23
II.4 Conclusion.....	24

Chapitre III: Simulation du tatouage audio par le masquage perceptuel

III.1 Introduction.....	25
III.2 Etapes de l’implémentation.....	26
III. 2.1 Mise en forme du signal à tatouer.....	26
III.3 Détermination du masque fréquentiel.....	27
III.4 Calcul du spectre de puissance.....	28
III.5 Identification des composantes tonales et non-tonales.....	29
III.6 Elimination des sons masqués.....	30
III.7 Détermination du seuil de masquage fréquentiel.....	30
III.8 Détermination du masque temporel.....	31
III.9 Génération de la signature de l’auteur.....	32
III.10 Masquage de la signature.....	33
III.11 Pondération avec le masque fréquentiel.....	33
III.12 Pondération avec le masque temporel.....	35
III.13 Tatouage d’une trame.....	35
III.14 Tatouage global.....	36
III.15 Extraction et vérification du tatouage.....	36
III.16 Signature représentée par un « Bruit Unique».....	37
III.17 Bruit unique généré aléatoirement.....	37
III.18 Bruit unique généré pseudo-aléatoirement.....	37

III.19 Signature représentée par une séquence audio.....	38
III.20 Conclusion.....	38

Chapitre IV : Critères et mesures de performances d'un système de tatouage

IV.1 Introduction.....	39
IV.2 Fidélité perceptuelle.....	39
IV. 2.1 Partie voisée.....	39
IV. 2.2 Partie contenant du bruit.....	40
IV. 2.3 Partie silencieuse.....	41
IV.3 Fiabilité des résultats.....	42
IV. 3.1 Cas sans transmission (local).....	42
IV. 3.2 Cas d'une transmission.....	43
IV.4 Robustesse.....	43
IV.5 Coût et débit.....	43
IV.44 Conclusion.....	44
 Conclusion générale.....	 45
 Bibliographie	 46

Introduction général :

L'avènement des documents sous format numérique et la facilité de duplication et de modification permettent une large prolifération des médias numériques (audio, vidéo et image). Toutefois, cette accessibilité accroît le problème associé à la protection des droits d'auteur. Pour cette raison, les créateurs et distributeurs des documents numériques hésitent à rendre accessible leur propriété intellectuelle et sont à la recherche d'une solution fiable au problème de la protection des droits d'auteur des données multimédia.

Le tatouage numérique a été proposé comme un moyen d'identification de l'auteur ou du distributeur des données numériques. Le tatouage est un processus d'insertion d'information concernant le droit d'auteur, en opérant un léger changement sur les échantillons de la donnée. Contrairement au cryptage, le tatouage ne restreint pas l'accès à la donnée, mais la protège tout au long de son utilisation. Le tatouage est inséré de telle sorte à résider en permanence dans la donnée hôte. Quand le propriétaire est mis en question, l'information peut être extraite de la donnée tatouée pour prouver la propriété unique.

Dans ce travail, nous mettons en place une solution pour la protection de la propriété intellectuelle des œuvres numériques audio. Notre solution procède par la modification directe des échantillons du signal audio, en ajoutant une marque indélébile et secrète.

Le défi majeur d'un système de tatouage audio est l'inaudibilité des informations insérées. Pour cela, nous utiliserons la propriété du masquage perceptuel qui exploite les imperfections du système auditif humain. Le masquage, dans le domaine fréquentiel, est le plus souvent utilisé dans les schémas de tatouage audio ; cependant, il ne procure pas toujours des résultats irréprochables. Par conséquent, nous introduisons en plus du masquage fréquentiel, le masquage dans le domaine temporel, ce qui donne un double masquage que subiront les informations insérées. En effet, le masquage temporel corrige certains défauts que peut engendrer le masquage fréquentiel sur le signal tatoué.

Chapitre I: Généralités sur le Tatouage audio

I.1 Introduction:

Le problème d'échange de données secrètes a toujours existé, et ce depuis la naissance des grandes civilisations. Les avancées technologiques en informatique et télécommunications ont contribué à soulever une multitude de problèmes liés à la protection (sécurité) de l'information, en permettant ainsi un large développement scientifique et technique en réponse aux défis soulevés. Parmi les questions posées, nous citons : la protection des droits d'auteurs, et la vérification de l'intégrité des données. A cela il est possible d'ajouter des questions telles que l'authentification, l'accès conditionnel, le tatouage numérique, la signature numérique, la communication secrète et la stéganographie. Notre système auditif néglige certaines parties de l'information contenue dans un signal audio. Cette propriété est utilisée pour tatouer des signaux numériques, c'est à dire pour y inclure une information supplémentaire non perceptible. Cette technique s'est beaucoup répandue, au cours des 10 dernières années, face à la forte demande industrielle en matière de protection de la propriété intellectuelle. Mais plus généralement elle permet d'ajouter dans la transmission des données audio un canal supplémentaire caché, le tatouage (ou Watermarking en anglais) est l'art de cacher de l'information directement dans des données multimédia de façon robuste et imperceptible. Dans le contexte des signaux audio, le tatouage met à profit les imperfections du système auditif humain pour garantir l'inaudibilité du message inséré [1].

Initialement, le tatouage audio s'est fortement développé avec l'augmentation de documents sous format numérique. Les signaux audio, auxquels nous nous intéressons, sous leur forme numérique, sont très facilement reproductibles. Ainsi, des techniques de protection efficaces, telles que le tatouage, sont donc devenues indispensables pour permettre d'authentifier les auteurs des documents, le tatouage consiste alors à insérer un signal (une marque ou signature) dans un autre signal numérique, audio dans notre contexte. Le signal tatoué résulte de la superposition de ces deux signaux : le document tatoué contient alors des informations (également appelées données cachées) qui peuvent être utilisées à plusieurs fins [2].

La figure suivant représente les déffirentes Techniques de la sécurité de l’information:

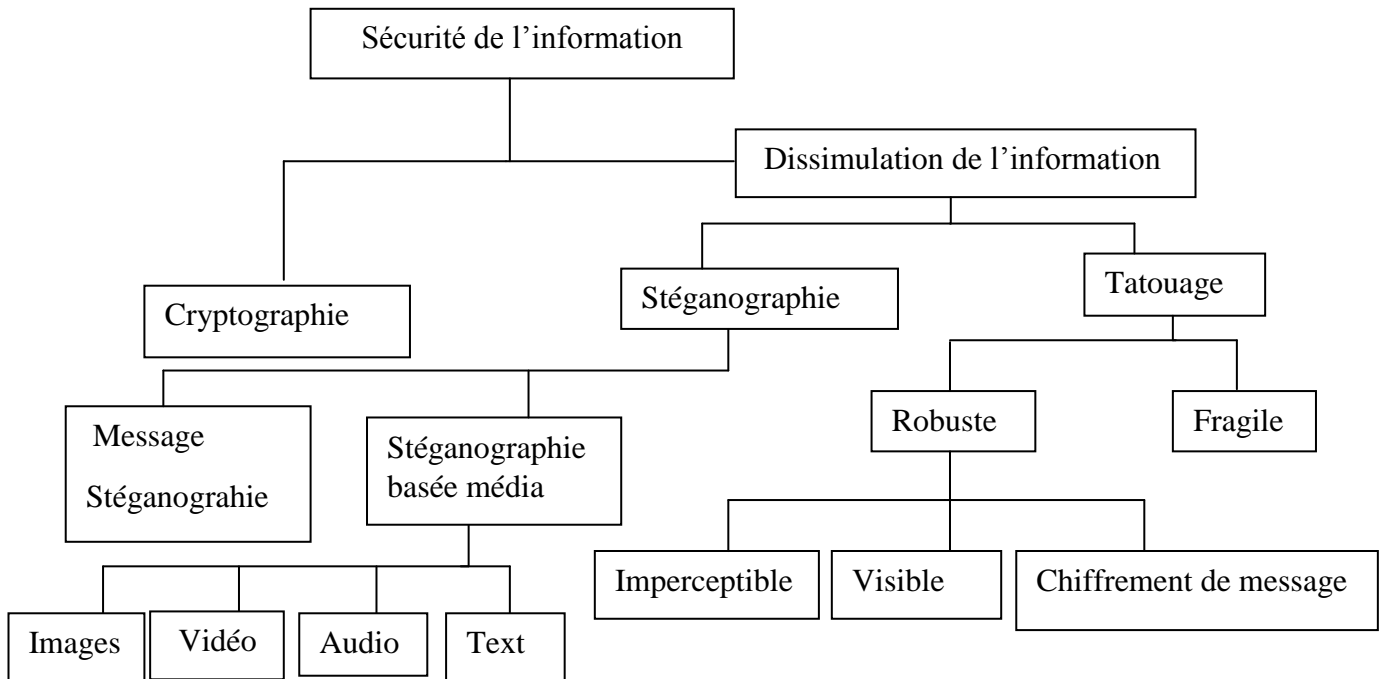


Figure 1.1: Techniques de la sécurité de l'information [3].

I.2 dissimulation de l’information, stéganographie et tatouage:

La dissimulation de l'information existe depuis plusieurs milliers d'années, rendre illisible un contenu en procédant à un cryptage n'est pas toujours la solution la plus adéquate en pratique. Dans certains environnements hostiles, cacher l'existence d'une communication est nécessaire pour éviter un certain nombre d'attaque de la part d'adversaires malintentionnés. Par exemple, la stéganographie qui fait partie de la dissimulation de l'information existait déjà en Grèce Antique. L'histoire de la dissimulation de l'information a souvent été liée aux applications militaires et au contre-espionnage. Durant les dernières décennies, la dissimulation de l'information a connu une grande révolution avec l'avènement d'Intente et des supports numériques. Ainsi, la communauté a vu la publication de centaines d'articles et de communication traitant des dièreses techniques d'utilisation de la dissimulation de l'information. Elle est devenue l'un des outils les plus efficaces dans la lutte contre le piratage multimédia et dans l'encadrement de la diffusion des données numériques [4].

I.2.1 La dissimulation d’information:

La dissimulation de l'information traite du problème de transmission d'un signal représentant souvent un message ou une donnée à transmettre en utilisant comme ou canal de communication un

contenu hôte. Ce dernier peut prendre plusieurs formes, numérique : vidéos et images, ou encore les lettres d'un texte anodin utilise par les services secrets des deuxièmes guerres mondiale donc La dissimulation de l'information consiste à transmettre un message m via un document hôte donne par le vecteur $s \in \mathbb{S}^N$ ou N est un entier naturel.

Le signal s est modifié de façon à permettre au décodeur d'extraire le message sans que les caractéristiques du signal hôte ne soient modifiées, autrement dit, l'insertion de l'information doit idéalement être invisible du point de vue statistique et perceptuel.

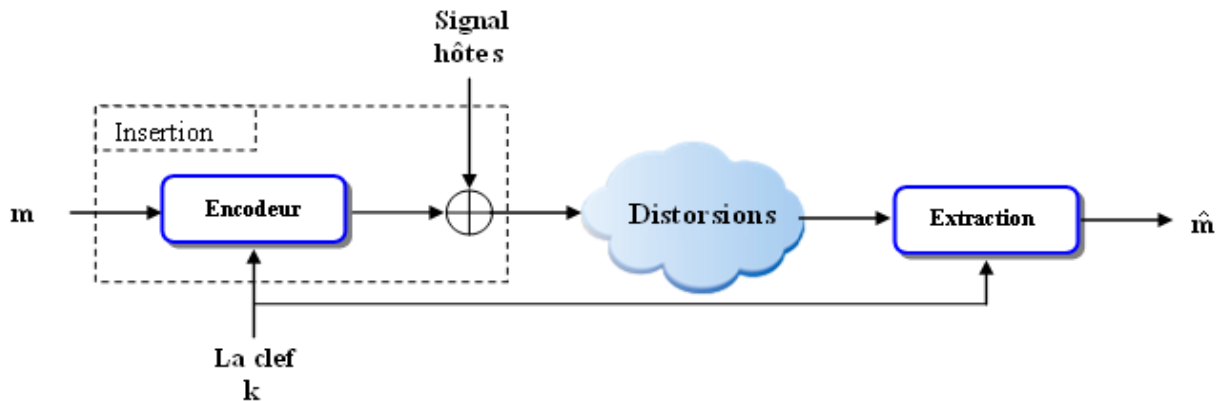


Figure 1.2: Schéma général d'un système de dissimulation d'information [4].

I.2.1.1 Caractéristiques d'un système de dissimulation d'information:

Dans Tous les schémas de dissimulation de l'information, il faut respecter un compromis entre leurs principales caractéristiques : Robustesse, capacité, invisibilité et sécurité. Ce compromis répond à certaines conditions dictées par un ensemble d'éléments : Le niveau de menace qui pèse sur l'information insérée, l'environnement dans lequel évolue le document marque, etc. Ces conditions sont liées aux types d'applications souhaites du système. Nous dévissons dans la suite les principales caractéristiques principales d'un système dissimulation d'information :

- ✓ **Capacité:** C'est le nombre moyen de bits transmissible par échantillon hôte. C'est également la taille du message qu'il est possible d'insérer sans erreur sur une longueur donnée du signal hôte pour une puissance d'attaque donnée. La capacité dépend principalement de la nature du signal hôte et du compromis robustesse/capacité.

Dans certains systèmes, la résistance au bruit est obtenue grâce à la redondance des bits d'informations, tel que le tatouage par étalement: En augmentant la redondance, la capacité diminue pour un gain en robustesse. Pour notre travail, la capacité du système de marquage n'est pas très importante puisque un identifiant comportera au maximum de 50 bits d'informations. Ce qui est faible relativement au support dans lequel l'indentant est insère.

✓ **Robustesse:** Ce terme est généralement réservé au cas spécifique de la résistance aux traitements légaux : compression, recadrage, changement du contraste...etc., ou l'attaque altère ou élimine la marque. Dans la protection de propriété intellectuelle et le traçage de document illicites, la robustesse prend une importance élevée puisque dans ce cas la marque doit pouvoir survivre aux différentes agressions ou toute modification du signal marque susceptible d'éliminer le tatouage sans qu'elle découle forcément d'une mauvaise intention de la part de l'utilisateur, par exemple le changement du contraste des séquences d'images ou la compression. La protection de notre document repose donc entièrement sur le tatouage, cette caractéristique est nécessaire pour garantir une diffusion sécurisée dans le cas du tatouage robuste et contrer ainsi les attaques du gardien actif en stéganographie.

✓ **Invisibilité:** Toute dissimulation d'information engendre inévitablement de modification sur le document hôte. Ceci impacte les caractéristiques perceptuelles et statistiques du signal original. Ainsi, le but de la majorité des systèmes de data hiding est d'impacter au minimum les caractéristiques du signal original.

✓ **Fidélité ou incivilité perceptuelle:** La distorsion due au marquage du signal hôte ne doit pas dépasser le seuil de perception de l'utilisateur, un nombre important de travaux ont été réalisés sur la modélisation du système sensoriel humain. Les contraintes liées à l'effet de la perception des de formations engendrées par l'insertion du message dépendent fortement de la nature du signal hôte (image, audio ou vidéo, par exemple). Ainsi, pour un signal hôte vidéo, il faut tenir compte d'une contrainte supplémentaire par rapport au signal audio et image. Puisque la dimension spatio-temporelle du signal limite les performances des systèmes de dissimulation. En effet une marque peut être invisible pour une séquence d'image isolée, mais le défilement de plusieurs séquences dièreses marquées au même endroit la rend visible par l'utilisateur.

✓ **Invisibilité statistique ou indétectabilité:** Certains systèmes de dissimulation de l'information permettent un marquage du contenu sans aucun impact visuel mais modifient la distribution de probabilité du signal hôte. Ceci peut être problématique surtout dans le cas de stéganographie, puisque l'indétectabilité est une condition. Dans le cas du tatouage numérique robuste, un attaquant pourrait localiser les parties tatouées du signal en procédant à une étude statistique du signal marque. Il pourra ainsi mener de puissantes attaques ciblées éliminera le tatouage, sans altérer la qualité du signal global.

✓ **Sécurité de la dissimulation de l'information:** La Sécurité d'un système de tatouage est dénie par la difficulté qu'aura un utilisateur malintentionné pour détecter/retrouver l'information insérée ou pour supprimer la marque et récupérer le signal hôte originale. La sécurité d'un système de dissimulation de l'information repose, en partie, sur une ou plusieurs clefs cryptographiques utilisées

à l'insertion et à l'extraction de l'information. Par exemple, dans certains schémas de tatouage, l'information est insérée sous la forme d'un signal produit par un générateur pseudo-aléatoire dont la combinaison d'entrée représente la clef secrète. Il existe deux niveaux de sécurité, dans le premier, un utilisateur non autorisé ne pourra pas détecter, décoder ou lire le message insère. Dans le second, le message inséré est crypté et ne peut être lu qu'à l'aide d'une clef secrète. Autrement dit, il faut que le système de dissimulation de l'information soit capable de contrer toute attaque venant d'une personne malintentionnée et la clef de cryptage doit rester la dernière défense contre les attaques [4].

I.2.2 La stéganographie:

La stéganographie (du grec steganos, couvert et graphein, écriture) est l'art de cacher un message secret au sein d'un autre message porteur (texte, image, son, vidéo...) de caractère anodin, de sorte que l'existence même du secret en soit dissimulée. Alors qu'avec la cryptographie, la sécurité repose sur le fait que le message chiffré soit incompréhensible pour les personnes non autorisées, avec la stéganographie, la sécurité repose sur le fait que la présence même d'un message secret ne sera sans doute pas soupçonnée et détectée [3].

Il existe deux types de stéganographie :

- La sténographie linguistique
- La stéganographie technique

La sténographie linguistique:

La littérature sur la stéganographie linguistique, dans laquelle les propriétés linguistiques d'un texte sont modifiées pour cacher l'information, est faible par rapport à d'autres médias

(Bergmair, 2007) [3]. La raison probable est qu'il est plus facile d'apporter des modifications aux médias non linguistiques dans lesquels le message secret sera indétectable par un observateur.

La stéganographie technique:

La stéganographie technique regroupe toutes les techniques qui ne jouent pas sur les mots, et elle est intéressante car elle permet de dissimuler des données dans plusieurs types de médias.

I.2.3 Tatouage ou Watermarking:

Le tatouage d'un signal (Watermarking) consiste à insérer et cache une information binaire dans un signal de façon imperceptible, il consiste aussi à insérer une signature ou information inaudible dans un signal audio. Cette information représente une séquence pseudo aléatoire. La détection du tatouage réalisée en utilisant une corrélation entre le tatouage original et le tatouage estime à partir du signal observe. Comme il introduit la notion de robustesse par rapport à la stéganographie [3].

I.2.3.1 Principe du tatouage:

Le principe du système est schématisé par la (Figure 1.3). On distingue trois principaux éléments de cette chaîne, qui seront détaillés par la suite :

-L'émetteur qui consiste en la mise en forme du message binaire a_k en un signal de tatouage $w(t)$ et à l'insertion par addition dans le domaine temporel de $w(t)$ dans le signal audio $x(t)$. Donc, on obtient le signal tatoué $y(t)$:

$$Y(t) = x(t) + w(t) \tag{1.1}$$

-Le canal de communication qui est à l'origine des perturbations apportées au signal audio tatoué, une compression audio à bas ou à très bas débit dans notre contexte.

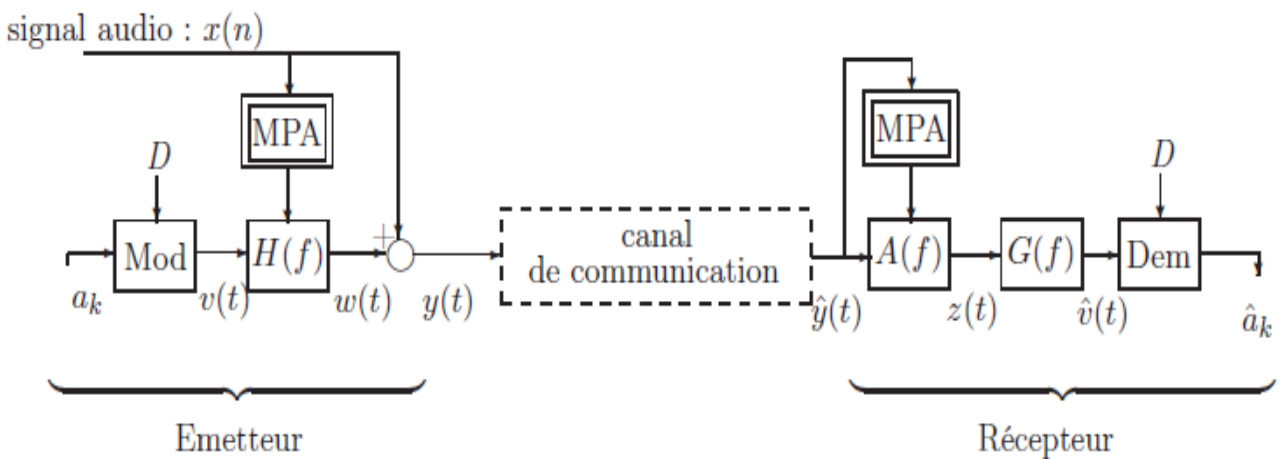


Figure 1.3: Système de tatouage de référence en présence de perturbations externes [5].

-Le récepteur qui doit restituer l'estimation \hat{a}_k du message binaire inséré à l'émission.

- **L'émetteur : générateur du signal tatoué:**

Le rôle de l'émetteur consiste à insérer le message binaire désiré dans le signal audio hôte. Le signal de tatouage doit être imperceptible et robuste aux différentes manipulations dont il pourrait être éventuellement l'objet. Comme indiqué sur la figure (1.2), l'émission consiste en deux opérations : une modulation et une mise en forme spectrale du signal de tatouage $w(t)$.

- **Le récepteur : égalisation, dé-bruitage et détection:**

En se référant aux résultats théoriques en communications numériques, dans le cas classique d'un canal Bruit Blanc Additif Gaussien (BBAG) et en absence d'interférences entre symboles (IES), le récepteur optimal est un détecteur recherchant dans D le mot le plus proche de celui reçu.

Si l'on considère le contexte de tatouage audio, le canal de communication introduit :

- Un bruit très fort, corrélé et non gaussien, constitué par le signal audio $x(t)$.
- Des IES dues au filtre de mise en forme H .
- Une non-stationnarité du canal liée à la variation d'une fenêtre à l'autre de la Réponse impulsionnelle (RI) de H .

Ces contraintes pèsent sur le choix de récepteur optimal à étudier l'apport des structures d'égalisation dans ce système de tatouage. Elle a montré en particulier que le problème posé par la chaîne de tatouage était équivalent à une égalisation aveugle d'un canal non causal et à minimum de phase [5].

I.3 tatouage numérique:

Le tatouage numérique est une discipline très récente. Sa naissance remonte au début des années 90 avec l'article de Tanaka. Digital Watermarking (tatouage numérique) fut pour la première fois employé en 1993 par Tirkel [6]. Le tatouage des signaux audio numériques, outre ses applications de protection de la propriété intellectuelle, révéla la possibilité d'utiliser le signal audio numérique comme un support de communication : une information peut être insérée imperceptiblement dans un signal audio et être ainsi diffusée en suivant le même canal de transmission que celui emprunté par le signal audio.

I.3.1 Définition:

Le tatouage numérique consiste à transmettre un message m via un support hôtes. Ainsi, le signal est modifié de façon à permettre au décodeur d'extraire le message sans que les caractéristiques du signal hôte soient modifiées, autrement dit, l'insertion de l'information doit être parfaitement imperceptible.

L'opération de tatouage peut être résumée dans le schéma suivant. La marque est choisie dans l'alphabet M , i.e., $m \in \{1, \dots, N\}$, ou N est un entier naturel.

Le signal tatoué $x = s + w_m$, ou w_m est le signal water mark obtenu à partir du message m , est transmis sur le canal ou il subit diverses distorsions et dégradations. Ensuite, le décodeur donne une estimation du message à partir du signal reçu.

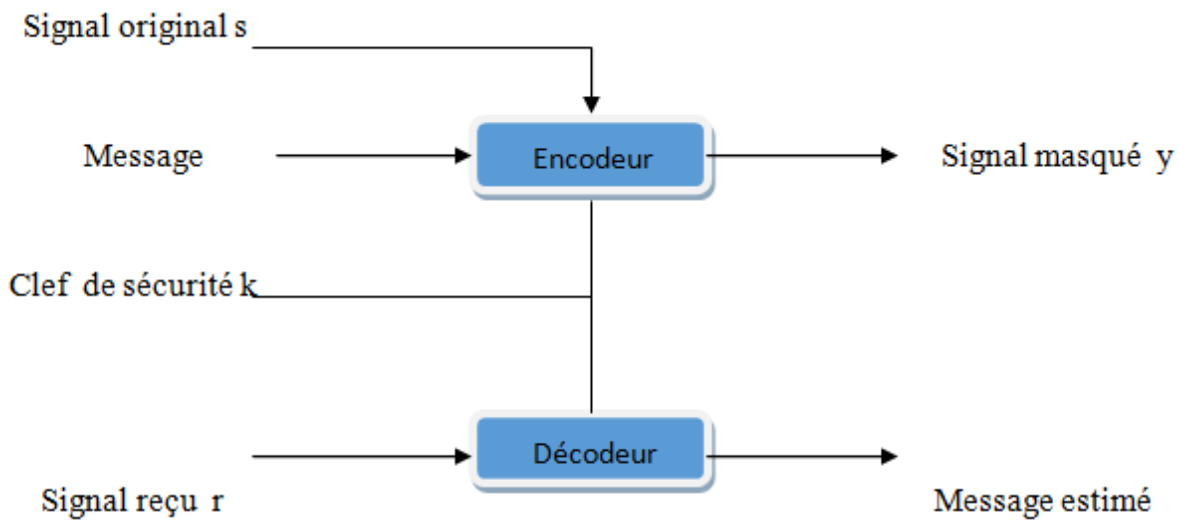


Figure 1.4 : Schéma Générale d'un système de tatouage numérique [4].

I.3.2 Caractéristiques du tatouage numérique:

Tout schéma de tatouage découle inévitablement d'un compromis entre les principales caractéristiques du tatouage : Robustesse, sécurité, capacité et transparence. Ce compromis répond à certaines conditions dictées par un ensemble d'éléments (le niveau de menace qui pesé sur notre marque, l'environnement dans lequel évolue le document marque, etc...) qui sont liés à l'application à laquelle est destiné le système.

Sécurité: On définit la sécurité d'un système de tatouage comme étant la discute qu'aura un utilisateur malintentionné à retrouver l'information insérée ou à supprimer la marque et récupérer le signal original. La sécurité du Watermarking repose, généralement, sur une ou plusieurs clefs cryptographiques qu'on utilise a l'insertion et a l'extraction de notre information. Par exemple, dans certains schémas de tatouage, l'information est insérée sous forme d'un signal produit par un

générateur pseudo aléatoire dont la combinaison d'entrée représente la clef secrète. Il existe deux niveaux de sécurité, dans le premier, un utilisateur non autorise ne pourra pas détecter, décoder ou lire le message insère. Dans le second, le message insère est crypte et ne peut pas être lu qu'à l'aide d'une clef tenue secrète. Autrement dit, il faut que le système de tatouage résiste à toute attaque venant d'une personne malintentionnée et la clef de cryptage doit rester une défense ultime seulement [6].

I.4 Tatouage audio:

Le système de tatouage peut être vu comme une chaîne de communication bruitée. En effet, le tatouage est le signal à transmettre et le signal audio est considéré comme un bruit. Le tatouage audio (ou audio Watermarking en anglais) est l'art de cacher de l'information directement dans des données multimédia de façon robuste et imperceptible. Dans le contexte des signaux audio, le tatouage met à profit les imperfections du système auditif humain pour garantir l'inaudibilité du message inséré. L'utilisation du tatouage audio comme canal auxiliaire de transmission pour véhiculer des informations additionnelles. Ces informations peuvent être destinées à l'auditeur (par exemple paroles d'une chanson, publicité ou informations sur le document) [7].

I.4.1 Les applications de tatouage audio:

- ✓ **L'annotation de documents sonores:** pouvant servir d'aide à l'indexation dans les bases de données : l'information cachée à destination de l'auditeur ou de l'administrateur de la base de données peut indiquer le nom de l'artiste, le lieu de l'enregistrement ou toute autre donnée relative au signal. Cette annotation peut être faite en studio ou en temps-réel lors.
- ✓ **L'amélioration de systèmes de transmission existants:** l'information peut par exemple permettre l'écoute stéréo dans un système de transmission initialement dédié aux signaux mono ou diffuser le signal audio sous format numérique caché dans la bande FM analogique.
- ✓ **L'identification du document:** Cette application, souvent associée au tatouage de protection, peut être envisagée dans un contexte de transmission de données existe lors que l'identifiant ne sert pas de preuve de propriété mais permet d'établir des statistiques sur l'utilisation du document. L'information identifie le signal audio lors de son transfert dans un réseau de diffusion (radio, télévision ou Internet). Elle permet, par exemple, de savoir sur quelle radio a été diffuse le signal audio, combien de fois, Ces statistiques peuvent intéresser notamment les analyseurs de marche. Cette identification peut également remplacer les systèmes de mesure d'audience (établis pour des transferts analogiques alors que les diffusions télévisées tendent à devenir numériques).

✓ **Le contrôle d'applications cible:** L'information peut être destinée à une application adjacente à l'écoute du signal audio, nécessitant la mise en œuvre d'un décodeur spécifique.

Deux exemples peuvent être cités. Le premier fait référence à certaines stations de radio FM, qui ont utilisé l'information cachée pour réduire le bruit lié à la transmission FM du signal audio. Le second est le projet RNRT ARTUS dont l'enjeu est d'utiliser le tatouage de la séquence audiovisuelle pour transmettre des informations de mouvement à un clone virtuel. Ce clone, synthétisé par un décodeur spécifique, sera incrusté dans l'image télévisée pour reproduire l'information du télétexte en langage des signes.

✓ **L'ajout d'informations publicitaires:** Le signal audio peut cacher des informations annexes à diffuser vers un auditeur. Un système de tatouage pour le télé-achat, l'IVDS a d'ailleurs été proposé : l'information cachée contient la référence du produit mis en vente par le télé-achat.

✓ **L'ajout de données supplémentaires au média :**

Plus récemment, les chercheurs se sont intéressés à utiliser le tatouage comme un moyen d'embarquer des informations supplémentaires sur le média, notamment pour l'indexation du contenu audio, pour faciliter l'utilisation et le stockage dans une base de données et inclure des métadonnées décrivant le contenu (ex : l'histoire ou les paroles d'une œuvre musicale) ou des informations à destination d'une application cible (ex : à destination des utilitaires de lecture).

Cette application peut aussi être utilisée en considérant le signal audio comme un support d'information, le système de tatouage se présente alors comme une chaîne de transmission bruitée, pour laquelle le tatouage est l'information utile et le signal numérique un bruit [1].

I.4.2 Les contraintes:

Les principales contraintes que doit satisfaire le système de tatouage varient selon les applications. Elles sont :

✓ **L'inaudibilité du tatouage :** Le signal de tatouage ne doit pas être perçu par l'auditeur afin de ne pas altérer la qualité sonore de la musique. L'inaudibilité de ce signal est assurée en exploitant les propriétés psycho acoustiques de l'oreille humaine et l'information tatouée ne doit pas dégrader perceptuellement le signal audio dans lequel elle est insérée : son insertion doit être transparente.

✓ **Le débit et la fiabilité de transmission :** La vocation du système étant de transmettre une information via un signal audio, elle doit permettre au même titre que les chaînes de communication classiques une fiabilité de transmission aussi élevée que possible pour un débit de transmission aussi grand que possible.

✓ **La robustesse:** Le signal audio tatoué peut être affecté par le canal de communication qui le véhicule, les performances de détection du message inséré peuvent donc se dégrader en présence

d'attaques de divers types, le tatouage doit résister à toute modification licite du signal (Compression, ré-échantillonnage) ainsi qu'aux perturbations illicites telles que le piratage.

✓ **Imperceptibilité:** le signal de tatouage doit être inaudible, autrement dit, il ne doit pas être perçu par l'auditeur.

✓ **Capacité :** c'est à dire la quantité maximale d'information pouvant être transmise pour une probabilité d'erreur quasi nulle, le débit de transmission doit être élevé [1].

Conclusion

Le tatouage est un procédé assez récent, inspiré de techniques plus anciennes. Dans ce chapitre nous avons introduit le sujet relatif à l'embarcation d'informations en général et au tatouage en particulier.

Chapitre II: Modèle psycho-acoustique et Masquage Perceptuel

II.1 Introduction:

L'acoustique est une science passionnante par son interaction avec une multiplicité de domaines très variés. Dans certains de ces domaines les problèmes posés sont si complexes que l'acoustique ne se développe véritablement avec profit que depuis quelques dizaines d'années. Cette complexité concerne aussi bien les phénomènes physiques, que les mécanismes du vivant, en particulier ceux de la perception auditive, qui ont évidemment une grande importance dans la vie de tous les jours. La branche de l'acoustique qui traite de la perception des sons s'appelle la psycho acoustique. Cette science étudie le fonctionnement de l'appareil auditif considéré comme une « boîte noire », en recherchant les relations entre les propriétés physiques des sons reçus et les sensations évoquées. Evidemment, elle est éclairée et complétée par l'étude des mécanismes mis en jeu, dans la physiologie et psychologie. Dans tous ces domaines il nous reste encore beaucoup à apprendre.

L'insertion du watermark consiste à ajouter à un signal audio un pseudo séquence aléatoire. Afin de garantir l'inaudibilité de la parole du watermark, la séquence doit être faite dans le domaine spectral selon un seuil de masquage obtenu utilisant un modèle psycho acoustique (Zwicker et Zwicker, 1991) [8].

Le masquage est les phénomènes où un son (masqué) est rendu inaudible en raison de la présence d'un autre son (appelé masquage) (peinter et Spanias, 2000) il est employé dans la Compression audio et le tatouage. Nous pouvons observer en particulier deux genres de masquage : Masquage simultané (fréquentiel) et masquage temporel (John, 1998) [8].

Dans l'algorithme de codage MPEG-1 l'allocation des bits est faite suivant un critère perceptuel. Elle est basée sur un modèle psycho acoustique dont le rôle est de déterminer une courbe de masquage et de fournir des rapports signal à masque à la procédure itérative d'allocation des bits. La résolution fréquentielle du banc de filtres d'analyse étant faible, les modèles psycho acoustiques proposés sont dérivés du spectre du signal de la trame obtenu par une transformée de Fourier discrète du signal temporel. Si un traitement est réalisé dans le domaine des sous-bandes sur le signal d'une trame qui sera soumis à une phase ultérieure de codage, une nouvelle allocation de bits doit être effectuée. Le problème essentiel est donc relatif à la détermination de cette allocation de bits. En effet, aucune information précise sur le spectre du signal ne peut être obtenue explicitement, sans recours au signal temporel, pour alimenter le modèle psycho acoustique.

II.2 Notion de psycho-acoustique:

I.2.1 Définition: La psycho-acoustique est l'étude des rapports entre les paramètres de la stimulation acoustique d'une part, et la qualité de la sensation auditive, d'autre part. Elle comprend l'étude de la représentation dans le système nerveux 'codage nerveux' des différentes dimensions des stimuli acoustiques, ainsi que l'étude du type et des règles de fonctionnement mis en jeu par les systèmes auditifs pour structurer la perception, y compris les mécanismes d'ordre supérieur appelés «processus cognitif»

La psycho-acoustique s'intéresse généralement à 4 types de perception : la *perception de l'intensité sonore*, la perception de la *hauteur tonale*, la détermination des *seuils*, l'audition binaurale et notamment la fonction de *localisation auditive*. La perception auditive a son niveau d'organisation de la perception de la parole et notamment la distinction du signal utile du reste de l'environnement sonore [9].

I.1.1.2 Modèle de Garcia: Généralement le modèle psycho acoustique de Garcia est une version du modèle psycho acoustique n°2 de MPEG (Motion Picture Expert Group). Garcia a proposé de calculer le seuil de masquage sur l'échelle des bandes critique pour respecter au mieux la physiologie de la perception auditive. La DSP du signal audio est transposée dans l'échelle de Bark par intégration de la puissance dans chaque sous bande.

La fonction d'étalement unique est appliquée pour la modélisation des effets de masquage de la membrane basilaire. Une translation vers le bas du seuil trouvé doit être faite avec un paramètre G coefficient du filtre $G(f)$ sous forme vectorielle. Ainsi, le choix optimal de ce paramètre de translation nous donnera le modèle le plus performant.

I.2.1.2 Modèle de Léandro:

Le modèle de Léandro est une version simplifiée du modèle n°1 de MPEG alors que la DSP du signal audio est décomposée en 4 sous bandes. En effet, dans chacune de ces sous bandes, la DSP subit une compression dynamique, une opération de convolution par unique fonction d'étalement (sans distinction de tonalité des composantes) puis un lissage. Le seuil obtenu est translaté par la suite pour obtenir un rapport de puissance entre le seuil de masquage et le signal audio (RMS) acceptable. Ce RMS influe sur les conditions d'in audibilité.

Cette translation (dans le domaine fréquentiel) est paramétrable, elle prend un facteur de gain α dans le domaine temporel. Ainsi, ce seuil offre l'avantage d'une très faible complexité de calcul (par rapport notamment au modèle MPEG).

La courbe du seuil de masquage est déplacée vers le bas d'un paramètre fixé a priori à -10 dB. Ce paramètre peut être réglé selon la puissance de tatouage souhaitée. Un choix optimal de ce paramètre nous garantira un modèle psycho acoustique plus performant [6].

I.2.2 Seuil d'audition absolu:

L'étendue des puissances acoustiques susceptibles de stimuler le nerf auditif sans douleur est vaste, allant de 0 à 120 dB SPL. Néanmoins, un son n'est détecté (et ne produit une résonance significative de la membrane basilaire) que si son niveau d'intensité est supérieure à un seuil fonction de la fréquence. Les expériences de Fletcher ont permis de quantifier ce seuil (sous la forme d'une quantité d'énergie moyenne en dB SPL) pour un son sinusoïdal pur de fréquence variable dans un environnement sonore non bruité. Cette quantification prend la forme d'une courbe appelée seuil d'audition absolu qui peut être approchée par la fonction non linéaire:

$$S_a(f) = 3.64f^{-0.8} - 6.5 \exp(-0.6(f-3.3)^2) + 10^{-3}f^4 \quad (\text{dB SPL}) \quad (2.1)$$

Les variations du seuil d'audition avec la fréquence constituent donc la courbe des seuils auditifs, qui se confond avec ce que nous appelons plus loin l'isotonique 0 dB. La forme générale de cette courbe s'explique par les caractéristiques des différentes parties de l'oreille:

- 1- l'ensemble pavillon et conduit auditif, en amont du tympan, est un canal de transmission acoustique qui possède comme on l'a vu une fréquence de résonance vers 3 kHz, et accroît globalement la sensibilité de 1 à 5 kHz.
2. les caractéristiques de transmission de l'oreille moyenne.
3. la répartition des neurones dans la cochlée : un plus grand nombre de neurones sont associés aux fréquences moyennes (50 à 5000 Hz), et la sensibilité globale à ces fréquences, liée au nombre moyen de décharges nerveuses, en est améliorée [1].

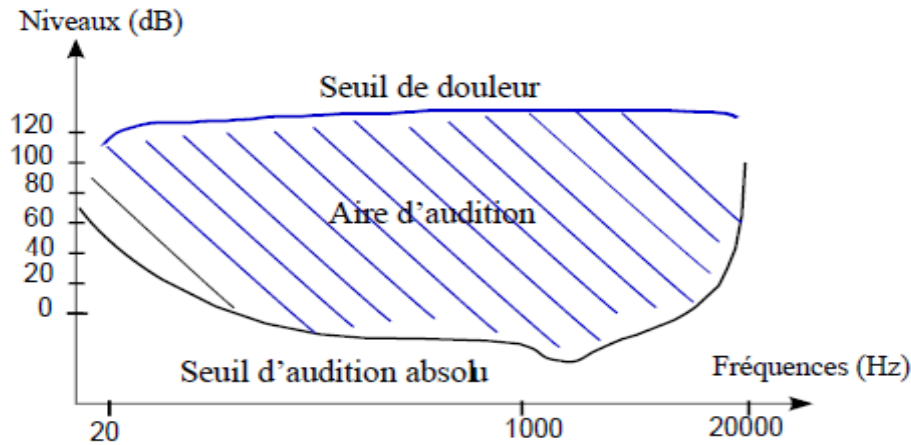


Figure 2.1: seuil audition absolu [10].

I.2.3 Bande critique:

La notion de bandes critiques est reliée à la manière dont l'oreille réalise la perception dans le domaine fréquentiel. Ainsi, les expériences ont montré que le système auditif se comporte comme un banc de filtres qui se chevauchent et dont les fréquences centrales s'échelonnent continûment. Il est alors modélisé par 25 filtres recouvrant la bande audible et dont la largeur de chacun correspond à une bande critique. Cette notion de bande critique peut être illustrée en considérant le niveau de puissance juste audible pour deux sons purs de fréquence différente et de même puissance. En effet, ce niveau reste le même pourvu que l'écart en fréquence ne dépasse pas une certaine limite qui est la largeur d'une bande critique, dès que cette limite est franchie ce niveau doit être doublé : la puissance perçue par l'oreille correspond donc à la somme des puissances sur chaque bande critique. Elle apparaît aussi dans le masquage fréquentiel, ainsi lorsqu'un son pur de fréquence f_0 est masqué par un bruit à large bande centré sur la même fréquence seules les composantes du bruit contenues dans la bande critique centré sur f_0 sont effectivement masquant et donc seule leur puissance intervient dans le calcul de la courbe de masquage.

La largeur des bandes critiques augmente avec la fréquence. Elle est égale à 100 Hz pour les fréquences centrales inférieures à 500Hz, au-delà, elle vaut environ 20 % de la fréquence centrale considérée. L'échelle de Bark introduite précédemment correspond tout simplement au numéro de la bande critique, le tableau suivant représente les bandes critiques avec ses fréquences et ses largeurs [11].

Tableau 2.1: Bandes fréquentielle critiques [10].

Bandes Critiques	Fréquences (Hz)			Bandes Critiques	Fréquences (Hz)		
	Basse	Haute	Largeur		Basse	Haute	Largeur
0	0	100	100	13	2000	2320	320
1	100	200	100	14	2320	2700	380
2	200	300	100	15	2700	3150	450
3	300	400	100	16	3150	3700	550
4	400	510	110	17	3700	4400	700
5	510	630	120	18	4400	5300	900
6	630	770	140	19	5300	6400	1100
7	770	920	150	20	6400	7700	1300
8	920	1080	160	21	7700	9500	1800
9	1080	1270	190	22	9500	12000	2500
10	1270	1480	210	23	12000	15500	3500
11	1480	1720	240	24	15500	22050	6550
12	1720	2000	280				

I.2.3.1 Echelle de Bark:

Sur la base des résultats précédents, Zwicker a déduit une échelle de fréquences proportionnelle à la largeur des bandes critiques, de la manière suivante (E. Zwicker et R. Feldtkeller : Psycho acoustique, Ed. Masson, 1981) :

L’ouïe peut former une bande critique en n’importe quel point de l’échelle des fréquences.

La largeur d’une bande critique, quelle que soit sa fréquence centrale, est appelée un Bark. Jusqu’à 1 kHz, cette largeur de bande est linéaire et est égale à 100 Hz. Au-delà de 1 kHz, elle varie dans une proportion logarithmique avec la fréquence [12].

En les rangeant arbitrairement l’une à côté de l’autre, on trouve dans la zone de fréquences de 20 Hz à 16 kHz, 24 bandes critiques l’indice de la bande critique dans laquelle se trouve une fréquence donnée, transformé en une variable continue, devient une nouvelle échelle de fréquences : le Bark (du nom de H. Barkhausen, qui s’est intéressé parmi les premiers à la mesure de la sonie).

A la suite de ces travaux il est courant encore d’entendre dire que notre perception décompose le son selon 24 bandes critiques. Rien n’est plus trompeur. Il faut comprendre, comme le note

Zwicker, que ces bandes critiques se « forment » en n'importe quel point, concrètement là où se trouve l'excitation sonore.

Quant au nombre de 24, il est comme le dit aussi Zwicker, tout à fait arbitraire, dans la mesure où il découle de la juxtaposition de bandes critiques dont les positions ne sont choisies qu'en fonction de leur largeur.

Cette juxtaposition suppose en outre l'hypothèse implicite, que les bandes critiques seraient comme autant de filtres passe-bande de formes parfaitement rectangulaires. Cette hypothèse simple ne correspond évidemment pas à la réalité.

I.2.4 La compression audio et psycho acoustique:

Le processus de la compression MP3 peut être décomposé en 3 étapes. D'abord, le jet audio d'entrée traverse batterie de filtres qui divise le signal en sous-bandes de fréquence.

Simultanément, il traverse un modèle psycho-acoustique qui utilise le concept de masque auditif pour déterminer qui ce peut ou ne pas être entendu dans chaque sous-bande. Le bloc d'attribution de bit réduit au minimum l'audibilité du bruit.

C'est généralement une attaque involontaire qui apparaît très souvent dans des applications de multimédia. Pratiquement toutes les données audio, vidéo et images qui sont actuellement distribuées par l'intermédiaire de l'Internet ont été compressées. C'est pour cette raison que le fichier multimédia doit avoir une robustesse face à la compression. La compression de donnée consiste à obtenir des fichiers plus légers, afin d'améliorer la vitesse de transfert sur Internet ou limité l'espace de stockage utilisé sur un disque dur. Il existe deux principaux types de compression :

✓ **La compression sans perte:** Appelée aussi « compactage », cette solution consiste simplement à coder les données binaires de manière plus concise dans un fichier. Elle permet ainsi de retrouver la totalité des informations après une procédure de décompactage.

✓ **La compression avec perte :** Concernant essentiellement les fichiers de média (image, son, vidéo), elle consiste en une « réduction » de l'information basée sur notre propre limite de perception. Puisque l'oreille n'est pas sensible à toutes les fréquences audio (seulement entre 20Hz et 20KHz), et que sa sensibilité varie selon certaines d'entre elle, il est possible de réduire la quantité de données de telle sorte que le résultat soit très ressemblant à l'original, voire identique, pour l'oreille humaine [13].

II.3 Masquage perceptuel:

I.3.1 La présentation:

Définition de masquage: On sait bien que l'oreille n'est capable de détecter un son pur que si son niveau de pression acoustique dépasse un certain seuil. Ce seuil dépend de la fréquence et la courbe exprimant cette relation s'appelle courbe de seuil d'audition absolu. Lors de l'écoute de deux sons purs de fréquences suffisamment proches cette courbe augmente, il apparaît un effet de masque : le son d'intensité plus forte rend totalement ou partiellement inaudible celui qui est plus faible. On parle alors de masquage fréquentiel(ou simultané). Pour un son masquant de fréquence f_0 et de puissance P_0 , on peut déterminer la courbe de masquage qui lui est relative en mesurant pour toutes les fréquences f de la bande audible [20 Hz, 20 kHz] le niveau de pression acoustique limite, d'un son à cette fréquence, pour qu'il soit masqué. Cette courbe dépend de la fréquence f_0 et varie selon la nature spectrale des sons masquant et masqué : modélisables par une sinusoïde ou plutôt par un bruit à bande étroite. Dans les différents cas de figure, ces courbes ont quasiment la même allure. Ainsi on montre que si on exprime la fréquence suivant une nouvelle échelle appelée échelle des Bark, ces courbes peuvent être modélisées par une fonction de forme triangulaire. Cette dernière s'appelle fonction d'étalement de la membrane basilaire. Elle présente deux caractéristiques importantes : la pente du côté des fréquences supérieures est moins aiguë que du côté des fréquences inférieures et elle dépend aussi du niveau du son masquant (elle est d'autant plus raide que celui-ci est faible). La valeur, $S_m(f_0)$, de la courbe de masquage en f_0 est inférieure à P_0 , la différence $P_0 - S_m(f_0)$ est appelée indice de masquage.

I.3.2 Masquage fréquentiel:

Le masquage simultané dépend des spectres d'amplitude du son masquant et du son Masqué. Le son est masqué si son intensité est inférieure à un seuil qui est déjà fixé d'avance. Les fréquences respectives des sons masque et masquant et de leurs natures (sinusoïdales ou bruitées). Ce phénomène peut être expliqué d'un point de vu anatomique. La présence du son masquant de forte intensité, provoque la vibration de la membrane basilaire. Cette vibration est relativement forte au point spécifique de la membrane correspondant à la fréquence du masquant et atténuée dans les zones avoisinantes. Une seconde vibration, liée au son masqué, ne peut donc être détectée que si elle dépasse la vibration (même atténuée) provoquée par le son masquant [1].

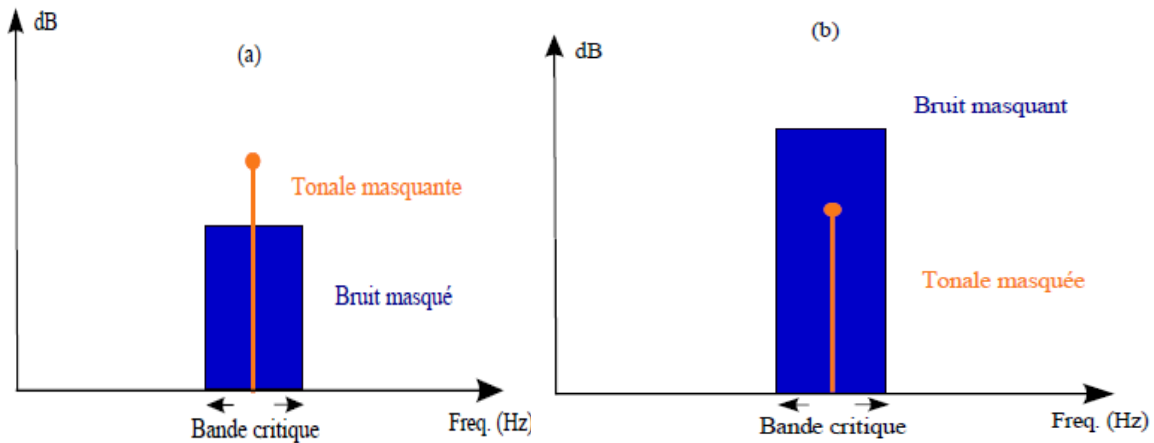


Figure2.2: Masquage fréquentielle [10].

I.3.2.1 Seuil de masquage:

Un seuil de masquage $SM(f)$ homogène a une densité spectrale de puissance qui établit la limite fréquentielle maximale admissible à satisfaire par le tatouage. Ce seuil de masquage, établi d’après un modelé psycho-acoustique, régit la mise en forme spectrale du vecteur V dont est issu le tatouage T . Cette mise en forme spectrale est réalisée soit dans le domaine temporel sous la forme d’une opération de filtrage, soit dans le domaine fréquentiel sous la forme d’une pondération fréquence par fréquence du tatouage [1].

Le seuil de masquage joue un rôle majeur dans la chaîne de tatouage : il garantit l’inaudibilité du signal de tatouage en présence du signal audio tout en permettant l’augmentation de la puissance du signal de tatouage par rapport à celle du signal audio, ce qui facilite sa détection. Le choix du seuil de masquage relève donc d’un compromis entre l’inaudibilité et le rapport de puissance entre tatouage et audio. A transparence du signal de tatouage équivalente, on aura donc tendance à choisir un seuil de masquage qui permet un rapport de puissance maximal. Le signal de tatouage étant construit pour que sa DSP coïncide avec le seuil de masquage `a un facteur multiplicatif α près, ce rapport peut être évalué par le RMS, défini par:

$$RMS = \frac{\int_0^{F_c} S_M(f)df}{\int_0^{F_c} S_X(f)df} \tag{2.2}$$

Ou $S_M(f)$ désigne le seuil de masquage et $S_X(f)$ désigne la DSP du signal audio $x(n)$. Ce rapport masque à signal est également un critère fréquemment utilisé pour caractériser la transparence du tatouage : plus le RMS est faible et plus le tatouage est inaudible. Le seuil de masquage doit également satisfaire `a une contrainte de robustesse liée au filtre de mise en forme perceptuelle. Les TEBs sont directement liés à la différence d’estimation entre le filtre de mise en forme calculé à

l'émetteur et sa version ré estimées à la réception : plus ces deux filtres sont semblables et meilleur est le taux d'erreur binaire. Cette conclusion se répercute directement sur les seuils de masquage, puisque les filtres sont construits de sorte à ce que leur réponse fréquentielle coïncide avec le seuil. Plus les seuils de masquage obtenus à partir du signal audio original et du signal audio tatoué sont similaires, meilleure en est la réception. Il convient donc de choisir un modèle psycho acoustique robuste, c'est à dire un modèle pour lequel le seuil de masquage $\hat{S}_M(f)$ calculé à partir du signal audio tatoué détérioré $\hat{y}(n)$ soit le plus similaire au seuil de masquage $S_M(f)$ du signal audio original:

$$\forall f, \hat{S}_M(f) \approx S_M(f). \quad (2.3)$$

I.3.2.2 Norme ISO/MPEG :

Le modelé psycho-acoustique de la couche 1 de la norme ISO/MPEG-1 est aussi conçu initialement pour déterminer la quantité maximale de bruit de quantification inaudible que peut introduire le processus de codage audio.

Les performances du système de tatouage et l'efficacité des stratégies d'insertion pour le cas d'une compression MPEG 1 couche 1 pour les deux débits de compression : 96 et 64 Kbits/s. Dans cette section, nous présentons une étude expérimentale des performances du système de tatouage dans le cas d'une compression MP3, Deux objectifs sont visés par cette étude. Le premier est de confirmer l'efficacité du système de tatouage en termes de Taux d'Erreur Binaire (TEB) en présence d'un canal de communication représenté par une compression MPEG (MP3) [10].

L'influence de la largeur de bande des vecteurs du dictionnaire d'émission sera également étudiée. Le deuxième objectif est d'évaluer les performances du système de tatouage de référence dans le contexte du multi-codage. Le but de la compression MPEG n'est pas d'obtenir un signal décodé égal au signal original, mais plutôt que ce signal soit perçu de la même manière que l'original par un auditeur humain. Ce principe est fondé sur propriété de masquage auditif. En effet, certaines composantes du signal sonore, particulièrement les parties hautes fréquences, ne sont pas perçues par l'oreille humaine en présence d'autres signaux. Ces parties masquées du signal sont donc considérées par le codeur comme insignifiantes. De ce fait, lors du procédé d'allocation binaire et sous la contrainte de débit de compression, peu de bits sont alloués à la quantification de ces composantes "presque" inaudibles et l'information qu'elles portent est, par conséquent, en grande partie perdue.

I.3.3 Masquage temporel:

On parle de masquage temporel lorsqu'un son de faible puissance est masqué par l'apparition à un instant différent d'un autre son de forte puissance, peut-être de deux natures, en fonction de l'ordre d'apparition des sons masquant et masqué. Si le son masqué est antérieur au son masquant, on parle de pré-masquage ; sinon, de post-masquage. Un son n'est masqué que si son intensité est inférieure à un certain seuil. Ce seuil est en fonction de l'écart temporel entre les deux sons, de la fréquence des deux sons, et de l'intensité et la durée du son masquant. Les phénomènes de masquage apparaissent dans le domaine temporel lors de fortes variations du signal, les signaux transitoires créant des zones de pré et de post-masquage importantes. Ainsi l'oreille ne perçoit pas les sons faibles précédant ou suivant immédiatement un son de forte intensité. Le post-masquage apparaît quand l'oreille a été stimulée par un son pur très fort a une certaine fréquence après cessation du son, il y a une perte de sensibilité autour de cette fréquence : environ 10 dB de perte qui disparaissent au bout de quelques centaines de millisecondes. Ceci traduit la persistance de l'effet masquant d'un son sur un autre. La durée effective du pré-masquage temporel est brève, de l'ordre de 5 ms, contrairement au masquage postérieur qui persiste plus de 100 ms (figure 2.3). Le phénomène de masquage temporel est difficile à modéliser et donc peu utilisé dans les applications audio [14].

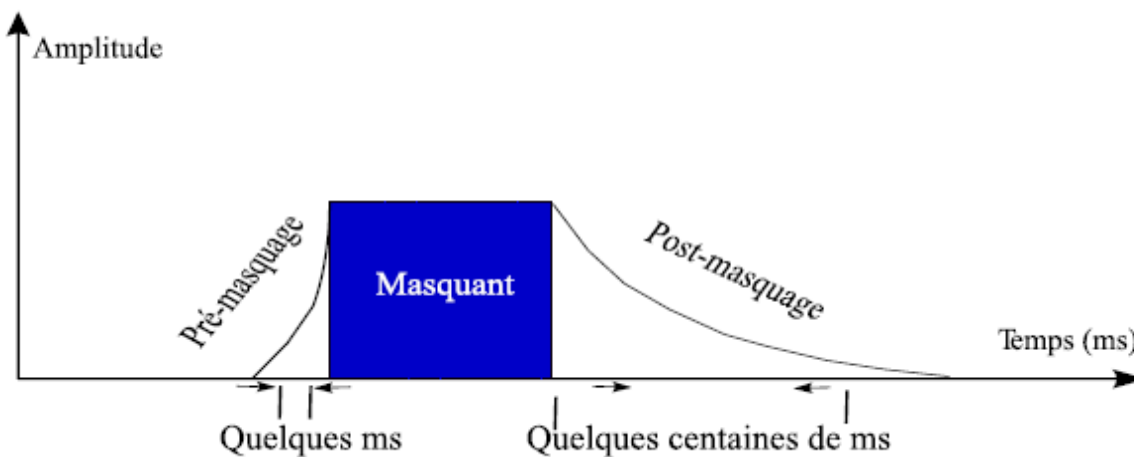


Figure 2.3: Masquage temporel [10].

I.3.3.1 La transformation de Hilbert:

David Hilbert (1862-1943) est né à Königsberg où il vécut, fit ses études et commença sa carrière jusqu'en 1895 où il s'installa à Göttingen. Ses recherches recouvrent un vaste domaine s'étendant de la théorie des nombres à la théorie de la démonstration en passant par la géométrie algébrique, le calcul vibrationnel et les équations intégrales. Ce sont ses travaux concernant le développement de fonctions arbitraires en séries de fonctions orthogonales [15].

Définition:

Le système linéaire dont la fonction de transfert.

$$G(u) = -\text{Sgn}(u) \tag{2.4}$$

A pour seul effet d'avancer les phases de $\pi/2$ est appelé un filtre en quadrature. La réponse impulsionnelle. $G(t)=1/\pi t$

Permet d'obtenir la réponse du système à une entrée $e(t)$.

$$S(t)=\frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{e(t)}{t-t} dt \tag{2.5}$$

Par définition $s(t)$ appelée la transformée de Hilbert de $e(t)$ est la notion suivante :

$$F(t)_{Hi} \equiv H\{f(t)\} \equiv \frac{1}{\pi t} * f(t). \tag{2.6}$$

La transformation de Hilbert intervient lorsqu'on étudie les signaux causaux.

Les signaux non stationnaires sont souvent étudiés via leur signal analytique lequel est calculé en utilisant la transformée de Hilbert. La transformée de Hilbert permet de calculer le signal analytique.

$$f_x(t)=f(t)+iH\{f(t)\}. \tag{2.7}$$

L'utilisation de la transformation de Hilbert dans l'étude des signaux causaux et des signaux non stationnaires sera vue dans des chapitres ultérieurs nous n'allons ici, ne discuter que quelques apparitions de la transformation de Hilbert en théorie du potentiel et en théorie ondulatoire.

Conclusion:

Dans ce deuxième chapitre nous avons donné premièrement une vue générale sur le model psycho-acoustique système auditif humain, pour comprendre les différentes transformations du signal audio pendant la transmission de l'information, puis nous avons présenté les phénomènes de masquage (temporel et fréquentiel), et après nous avons exposé la méthode de compression audio « MPEG » qui est couvrent un très grand nombre d'applications.

Chapitre III: Simulation du tatouage audio par le masquage perceptuel

III.1 Introduction:

Le tatouage de signaux à été dans un premier temps présenté comme une technique permettant de répondre au problème de la protection de droits d'auteur. L'idée de base consiste à ajouter une marque (encore appelée signature) aux données à protéger. La marque doit être imperceptible : dans le cas du tatouage audio la présence de la marque ne doit pas introduire de distorsion audible, et elle doit aussi être robuste vis-à-vis aux traitements habituels du signal ainsi qu'aux attaques malveillantes. Cependant, cette marque doit être détectable par les personnes autorisées. En plus de son utilisation pour la protection des droits d'auteur, les techniques de tatouage peuvent être envisagées pour la transmission de données superposées à un signal audio. On parle alors de "contenu augmenté". Plusieurs applications sont possibles dans ce contexte. Par exemple, si l'information prioritaire à transmettre est de la parole, il apparaît alors utile de transmettre des informations liées à une analyse audio.

L'objectif de ce travail est de présenter un schéma de tatouage d'audio applicable dans un contexte de communications numériques. Le cadre de ce travail nécessite des

Contraintes précises :

- Le débit de transmission du message inséré doit être le plus haut possible.
- Le système de tatouage doit également être robuste face à la compression de données, en particulier à la compression de type MPEG qui est souvent utilisée.
- Les attaques destinées à enlever la marque dans le signal de parole (piratage) ne seront pas prises en compte. Ceci car les signaux à transmettre ne sont pas des données sensibles.

D'après une étude détaillée sur le tatouage, nous avons pu tirer que le domaine fréquentiel est un bon espace de point de vue robuste et inaudibilité d'où l'idée d'utiliser la FFT et IFFT pour passer du domaine temporel au domaine fréquentiel et la transformation la plus utilisée dans le codage audio vue qu'elle permet une résolution fréquentielle plus fine tout en étudiant les effets de bordure. D'autre part, pour diminuer l'audibilité de la marque lors de l'insertion des bits de tatouage, nous avons exploité les propriétés du modèle psycho acoustique.

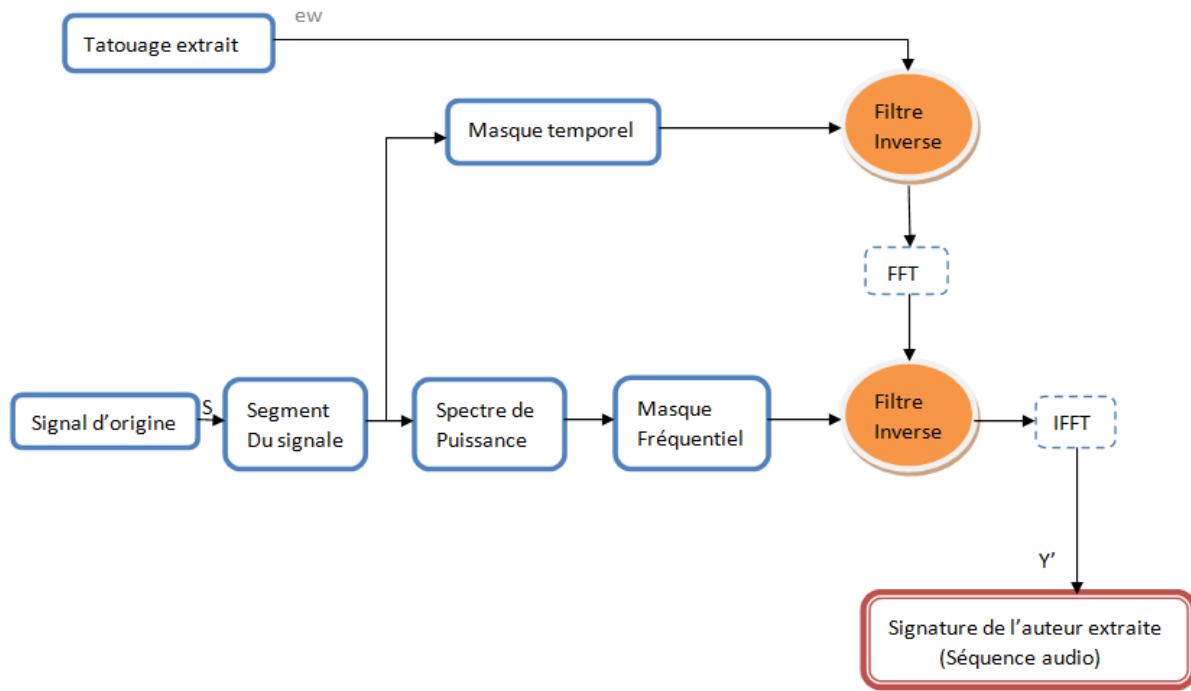


Figure3.1: Schéma général du tatouage.

III.2 Etapes de l'implémentation:

III.2.1 Mise en forme du signal à tatoué:

L'inaudibilité du signal de tatouage $w(t)$ est assurée par la mise en forme spectrale du signal modulé à l'aide d'un filtre perceptif $H(f)$ fondé sur un Modèle Psycho acoustique (MPA). Ce modèle, exploite les propriétés perceptives du système auditif humain. Cette opération vise à donner la puissance maximale à ce signal modulé tout en respectant les contraintes d'inaudibilité, la puissance du signal modulé ainsi maximisée optimise la détection à la réception. [1]

Le tatouage $w(t)$ est inaudible si sa Densité Spectrale de Puissance (DSP) est au dessous du seuil de masquage (fourni par le MPA) du signal audio sur chaque fenêtre de traitement.

La mise en forme du signal à tatouer fait la diminution du volume des fichiers audio (et donc du débit nécessaire pour leur lecture), une première étape dans la chaîne de compression consiste donc à éliminer tous les signaux dont l'amplitude se situe en dessous du seuil de perception.

Une séquence pseudo-aléatoire (la clé privée) est ajoutée au signal audio. Afin de garantir l'inaudibilité du tatouage, la séquence est mise en forme spectralement suivant un seuil de masquage obtenu à l'aide d'un modèle psycho acoustique. La détection est réalisée par une mesure de corrélation entre le tatouage original et un tatouage estimé à partir du signal observé. Cette détection est évidemment privée. [13]

III.3 Détermination du masque fréquentielle:

Ce phénomène étudié notamment par Fletcher et Zwitcker, met en évidence l'effet suivant: un son pur relève le seuil d'audibilité ou voisinage de sa fréquence. Il masque les sons d'intensité plus faible, situés à des fréquences proche de la sienne: voir la figure (3.2). On ne parlera plus alors de seuil d'audibilité, mais de seuil de masquage.

Les sons masqués constituent donc l'information non-pertinente, ils peuvent être éliminés du flux audio.

Cette propriété n'est pas typique des sons purs, mais elle s'applique également à un son quelconque, à condition de découper son spectre en bande fréquentielle et d'analyser le masque autour de chaque bande.

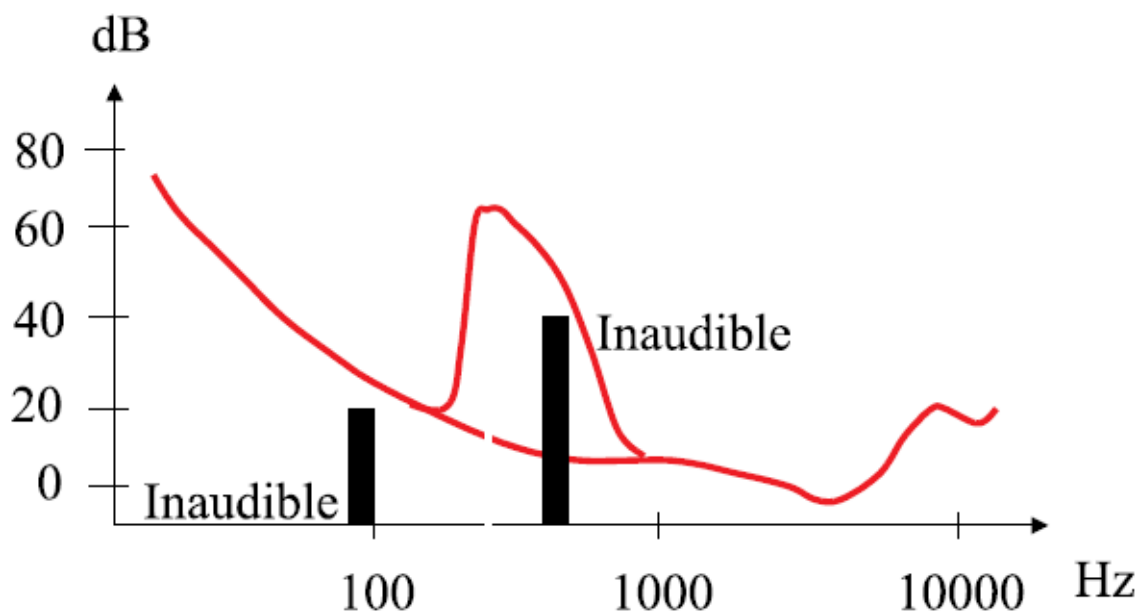


Figure3.2: masque fréquentielle [17].

Fletcher, puis zwitcker en me en évidence l'existence de bande fréquentielle critique appeler Barks ils en ont dénombré 24 sur l'ensemble du domaine audible. Tous les sons émis dans une même bande critique contribuent dans la même manière au masquage des bandes critiques voisines.

Par conséquent un bon modèle psycho acoustique devrait commencer par une décomposition du spectre en bande critique de Barks. Cependant, ceci compliquerait notablement l'implémentation, car ces bandes critiques n'ont pas des largeurs identiques, ni en valeur absolue, ni même en valeur relative, en moyenne et en haut fréquence les bandes critique sont fort mais pas en basse fréquence [17].

III.4 Calcule de spectre de puissance:

Chaque segment de 16ms du signal $x(n)$, $N=512$ échantillons est pondéré avec une fenêtre de hamming window, $h(n)$:

$$h(n) = \frac{\sqrt{8/3}}{2} \left[1 - \cos\left(2\pi \frac{n}{N}\right) \right] \tag{3.1}$$

Le spectre de puissance du signal est calculé:

$$S(K) = 10 * \log_{10} \left[\frac{1}{N} \left\| \sum_{n=0}^{N-1} S(n) h(n) \exp(-j2\pi \frac{nK}{N}) \right\|^2 \right] \tag{3.2}$$

Le maximum est normalisé à un niveau de pression acoustique de référence de 96 dB, le spectre de puissance du signal de test de 32 KHz est représenté en figure (3.3).

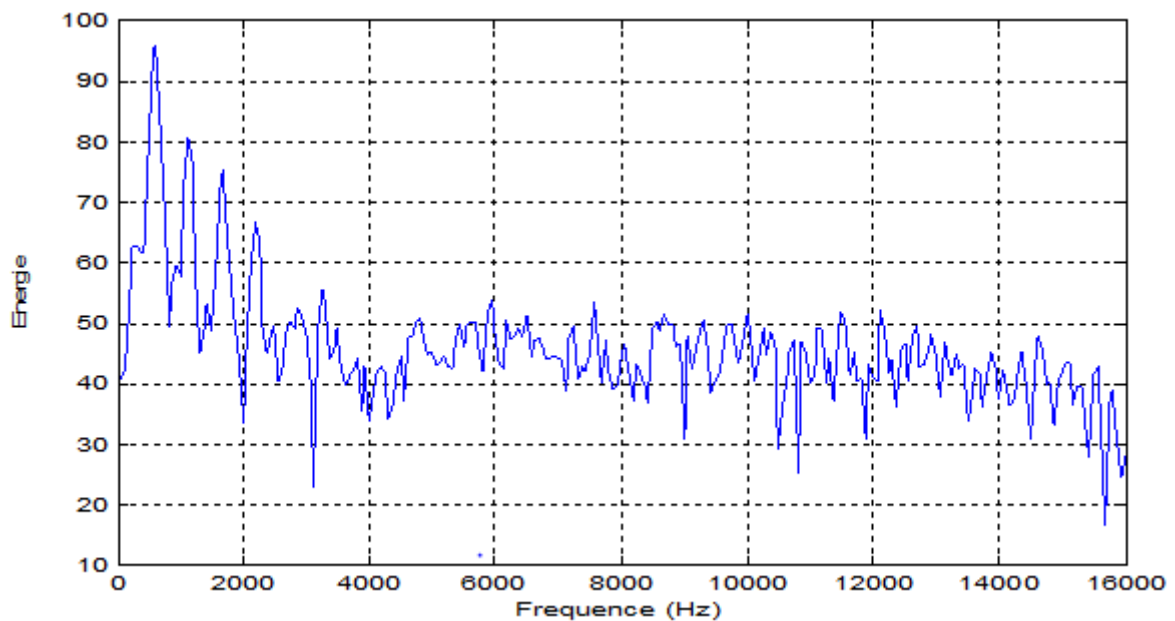


figure3.3: Spectre de puissance du signal.

III.5 Identification des composantes tonale et no tonale:

Dans le modèle MPEG, l'identification des composantes tonales et non tonales est un processus sélectif en fréquence. Un son pur, désigné dans le jargon de la psycho acoustique par le terme de tonale, généré une pression acoustique sinusoïdale dans le temps. Le niveau acoustique de ce son est représenté sur une échelle logarithmique. Il a l'allure d'une seule raie en fréquence. Un son pur est rarement rencontré dans la nature, une tonale au centre d'une bande critique va masquer tout bruit présent dans la même bande et d'intensité plus faible.

Les composantes tonales sont définies comme les maximums locaux du spectre de puissance c'est la composante est une tonale si elle est supérieure à ses voisines immédiates (maximum local). [9]

Non ajoutons à son intensité celles du composant précédent et suivant d'autres composants tonal dans la même bande de fréquence ne sont plus considérés. La composante non tonal est constituée de la somme des intensités de la composante du signal restant chacune des 24 bande critiques entre 0 et 15500Hz. Le système auditif comporte une bande de filtres passe-bande avec des fréquences centrales continuellement ouvertes ce filtre auditif peut être approché par des filtres rectangulaires avec une fréquence.

Dans ce modèle, la bande sonore est par conséquent divisée en 24 bandes critiques non régulières, les composantes tonales et non tonales du signal audio d'exemple sont représentées sur la figure (3.4).

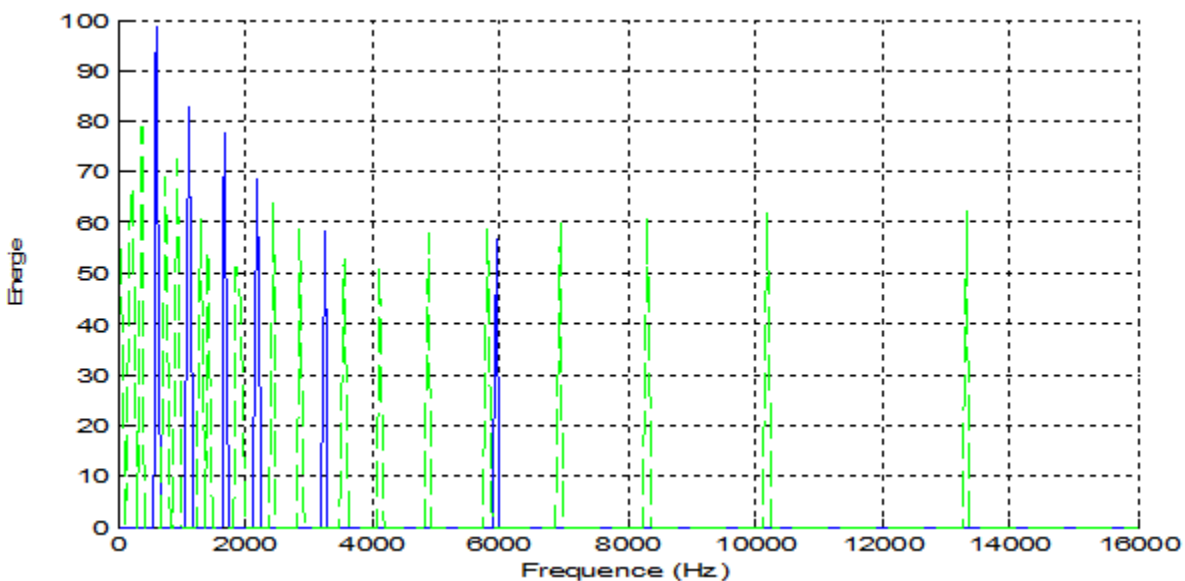


Figure3.4: les composants tonals.

III.6 Elimination des sons masque:

Dans la télécommunication et pour transmis les donnes entre l'utilisateur ce dernier fait un zone des fréquences bien déterminer pour éliminer tous les fréquences qui sont supérieur à un référence programmable appeler seuil de masquage à l'ide d'un threshold : le seuil (en dB sur le signal d'entrée) au-dessus duquel le signal est compressé.

Le composant doit être faible, le seuil d'audition absolue et le composant tonal séparé moins de 0.5 barks sont élimination (le bruit), et un complot des composants enlevés un long avec le seuil d'audition absolue est montré dans la figure (3.5).

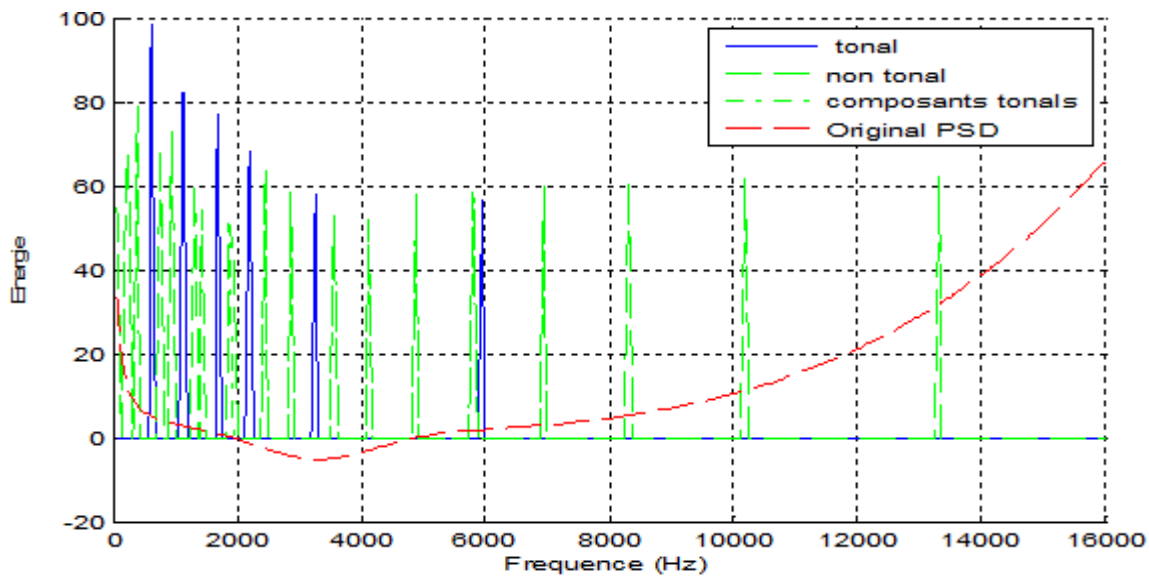


Figure3.5: élimination des composants masqués.

III.7 Détermination du seuil de masque fréquentiel:

Ce seuil représente la limite supérieure du bruit pouvant être ajouté au signal original par les opérations de quantification, sans que ce bruit devienne audible, elle présente par la relation de:

$$ATH(dB) = 3,64 * \left(\frac{f}{1000}\right)^{-0,8} - 6,5 * e^{(-0,6 * \left(\frac{f}{1000} - 3,3\right))^2} + 0.001 * \left(\frac{f}{1000}\right)^4 \quad (3.3)$$

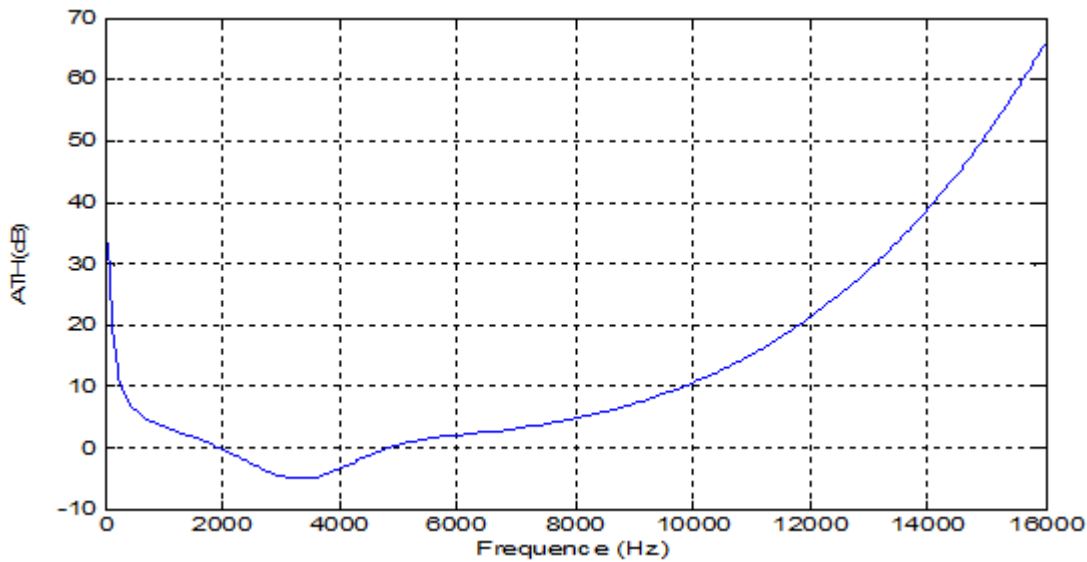


Figure3.6: seuil audition fréquentielle.

III.8 Détermination du masque temporel:

On parle de masque temporel lorsqu’un son de faible puissance est masqué par l’apparition à un instant différent d’un autre son de forte puissance alors d’après les études qui nous allons voir on peut expliquer ce phénomène: un son de forte intensité masque les sons de intensité plus faible, immédiatement postérieure (post masking) et antérieure (pre masking), la figure (3.6) montre que les propriétés sont ici asymétriques: le pré-masquage concerne un intervalle de temps de quelques millisecondes seulement, alors que le post-masquage s’étend dans plusieurs dizaines de millisecondes.

Le phénomène de pré-masquage semble à première vue contredire le principe de causalité, cependant, il ne faut pas oublier que notre oreille a un temps de réponse non nul, et qu’elle intègre, donc le stimulus sur une période correspondant à quelques 30ms. [4]

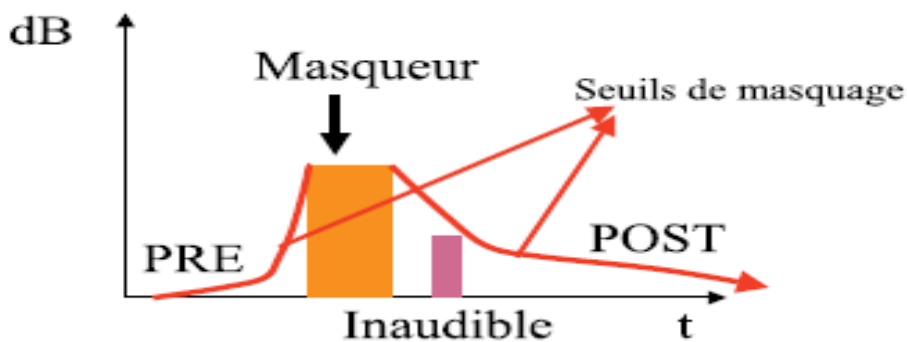


Figure3.7: masque temporel [4].

III.9 Génération de la signature de l’auteur:

Le principe de cette technique est d’insérer directement dans un signal audio, une marque voire une signature, contenant des informations particulières comme le titre d’une chanson ou le nom d’un artiste, alors la signature de l’auteur c’est une séquence binaire qui est ajouté par l’utilisateur au signal audio comme un bruit aléatoire, elle prendre la même forme de signal audio transmis mais avec un principe c’est qu’à une faible puissance et faible amplitude para pore au signal original.

On souhaite insérer la signature suivante:

$$Br = \text{bruit} (N/2) \tag{3.4}$$

Cette signature est tout d’abord sur-échantillonnée avec un facteur d’équation global afin de créer un signal redondant, permettant de le rendre plus robuste avec cette multiplication:

$$m = \text{treshs} * Br \tag{3.5}$$

Après la transformer de Fourier inverse avec la valeur absolu de cette signature on modulé ensuite le signal obtenu par un bruit binaire pseudo-aléatoire, qu’on amplifie ensuite par le multiplexage avec le filtre actife.

$$B = |ifft(m)| \tag{3.6}$$

En revienne à l’insertion de la signature, et le tatouage est ajouté de la façon suivante:

$$x_2 = x_1 \left(1 : \frac{N}{2} \right) + B \tag{3.7}$$

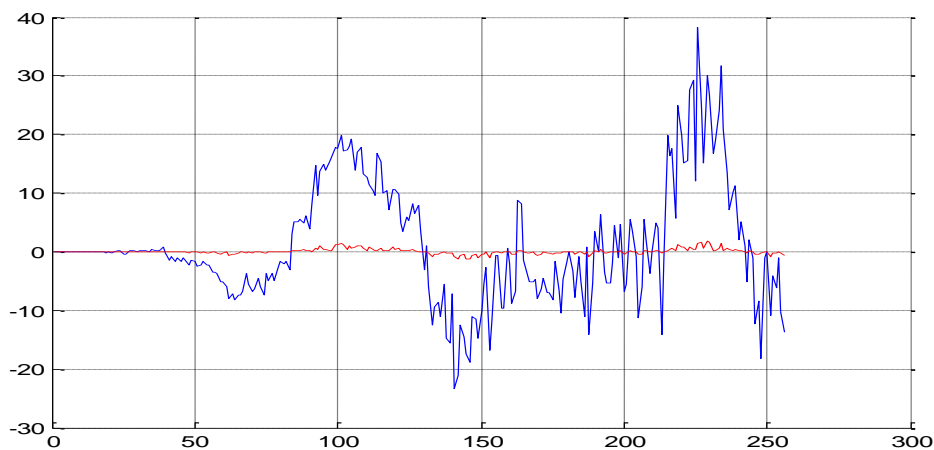


Figure3.8: Principe de signature para pore au signal audio.

III.10 Masquage de la signature:

Cette signature est masquée à l'aide du seuil de masquage figure (3.6), pour extraire l'information non-pertinente, les échantillons dont l'intensité est située sous le seuil de masquage seront éliminés, de plus, les échantillons restant seront en codés avec moins 16bits.

En codés avec moins de bits, cela revient à augmenter le bruit. Si l'on se rappelle l'analyse de rapport signal à bruit d'un encodeur. On tire de ce qui précède le principe de l'allocation des bits disponibles: le nombre de bit par échantillon doit être choisi de telle manière que le bruit reste inférieur de seuil de masquage (inaudible).

Un autre principe pouvant être utilisé pour réduire encore le nombre de bit est le suivant: lors le passage du signal audio le niveau de bruit (inaudible, masqué), n'est pas élevé, et donc la réduction du nombre de bit ne sera pas très importante.

On peut alors amplifier leur de l'encodage les niveaux du signal les plus faibles, ce qui permettra d'accepter un bruit, les niveaux du signal élevé reste eux inchangés: on à donc opéré une compression de la dynamique du signal.

Lors du décodage, l'opération inverse sera effectuée (expansion de dynamique): les niveaux les plus faible seront atténués du même facteur que lors de la compression, le même que le bruit qui passera sous le seuil de masquage [15].

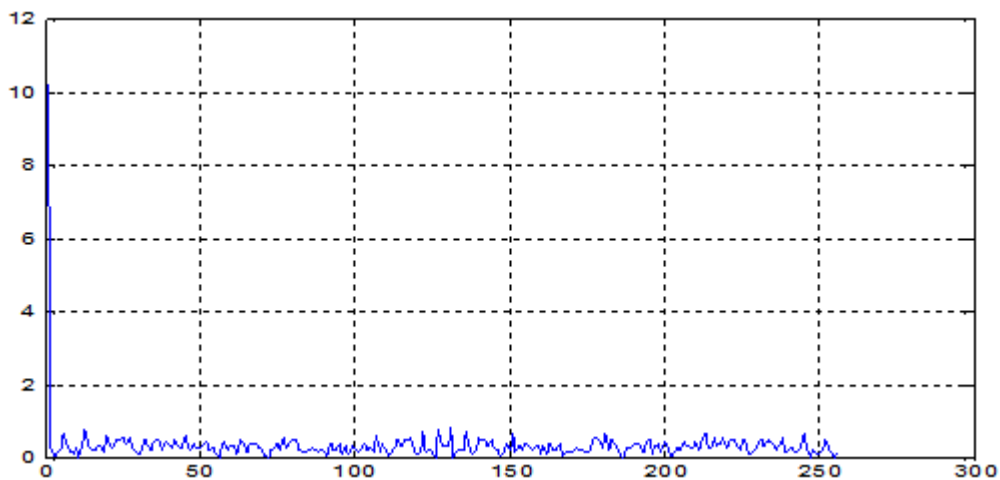


Figure3.9: bruit pondéré par le masque fréquentiel.

III.11 Pondération avec le masque fréquentiel:

L'analyse du phénomène de masquage lisse clairement sous-entendre qu'une analyse fréquentielle du signal est nécessaire, en particulier les premiers essais de réduire de bruit dans les années 1960 ont montré que la compression large bande était de mauvaise qualité, cela suggère qu'adopter un bruit.

Par conséquent il est nécessaire, de travailler par bandes de fréquence. Mais avec quelles largeurs de bande.

Nous avons déjà mentionné précédemment l'existence de bande critique (ou barks): le niveau de bruit masqué dépend uniquement de l'énergie du signal dans une bande critique, et non de sa répartition spectral dans la bande.

Une analyse par bandes critiques compliquerait cependant de traitement du signal, de plus elle n'est pas absolument nécessaire. Plusieurs codes perceptuels découpent le spectre audio en bandes de largeurs égales.

Le filtrage en sous bande sera effectuée par un bande de filtre, le spectre complet du signal est découpé en intervalles de largeurs égales, ils correspondent à ce découpage de filtre de type passe-bande, dans un premier temps, la sortie de ces filtres est un signal (filtré) échantillonné à la fréquence originale f_s .

Pour la première sous-bande (celle qui va de $L/N+1$), la décimation ou réduction de f_s ne modifié pas le signal, puisque les fréquences utiliser sont inférieur à la nouvelle fréquence de Nyquist, mais pour les autre sous-bandes, le critère de Nyquist n'est plus vérifié: par conséquence il y à recouvrement.

Ce recouvrement va néanmoins s'opérer de manière "propre" de telle manière que le signal original puisse finalement être récupérer.

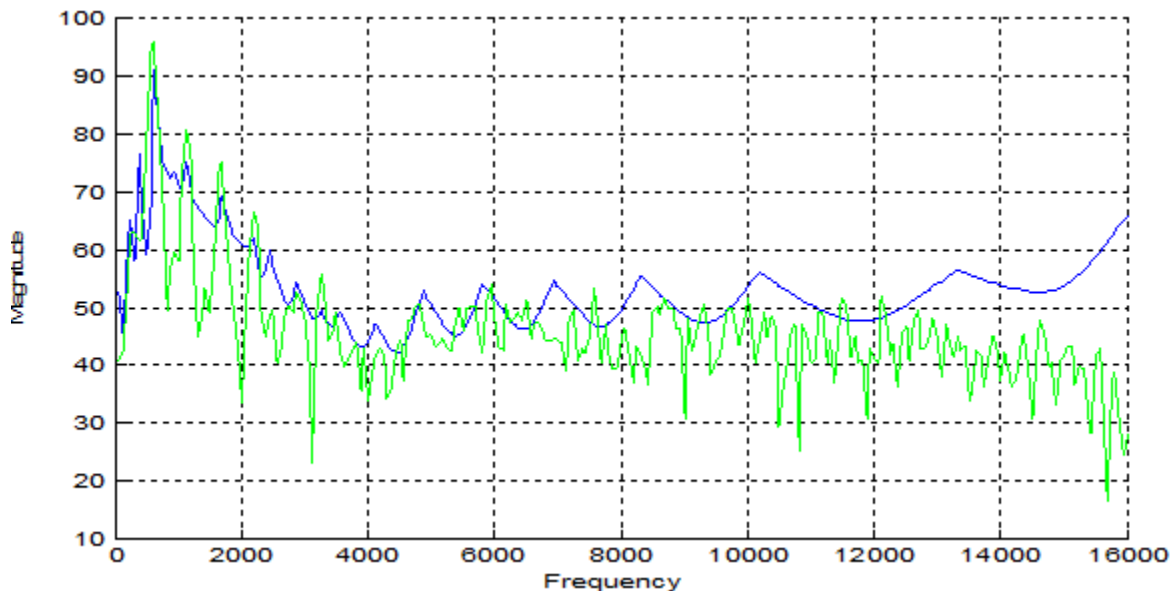


Figure3.10: masque fréquentiel.

III.12 Pondération avec le masque temporel:

Dans ce contexte, il reste à analyser le problème particulier de masquage temporel nous venons de voir que le niveau de bruit est constant dans une sous bande et dans un bloc. Il est déterminé par le nombre de bits attribués à la mantisse, est donc par "hauteur" de seuil de masquage dans cette région de spectre.

Remarque: pour le recouvrement des images de spectre s'effectue "proprement" les filtrages passe-bande devraient être idéaux. Même s'ils sont opérés dans le domaine numérique, un certain recouvrement subsiste malgré tout.

A la sortie du banc de filtres, on retrouve les signaux dont l'énergie est continue dans une sous-bande bien spécifique. Dans chaque sous-bande, il va falloir à présent estimer quel est le bruit acceptable, ce qui permettra de fixer le nombre de bits minimum à allouer aux échantillons du signal filtré.

L'estimation de bruit passe par la détermination du seuil masquage.

Pour déterminer le seuil de masquage, l'analyse de sous-bandes ne convient pas, car elles sont des largeurs constantes. Elles sont donc trop étendues en B.F. et trop étroites en H.F., par rapport aux largeurs de bande critique. Par conséquent, le seuil de masquage sera avantageusement évalué avec une résolution plus importante, par FFT (Fast-Fourier Transform) de 512 points.

Ainsi, la courbe continue de signal elle est évaluée à partir de la transformée de Fourier de ce signal, fréquence par fréquence, de la théorie du masquage fréquentiel cette transformée de Fourier est réévaluée périodiquement, en fait pour chaque "bloc temporel" du signal original.

On le voit, le codage perceptuel est basé sur une analyse temps-fréquence du signal audio: analyse par sous-bandes fréquentielles et par "blocs temporels". Les deux types d'analyses apportent leur lot de renseignements, à partir des quels le système de compression réussira à extraire l'information non-pertinente.

III.13 Tatouage d'une trame:

Le tatouage d'une trame dans un signal audio, c'est à dire la décision de classifier une trame en trame tatouée ou pas, peut passer par l'analyse trame à trame de l'énergie. Un changement soudain dans l'énergie ainsi qu'un changement soudain du contenu spectral peut s'apparenter à l'introduction d'une transitoire dans la trame, après cette technique on peut tatouer l'information inaudible à partir d'ajouter la signature à une partie de signal uniquement.

III.14 Tatouage global:

Dans ce type de tatouage la séquence binaire d'ajouté par l'auteur ou signal audio transmis ou-bien la signature est sa pendant tout la période de signal, sans modifier ou toucher la forme du signal original comme vu dans la figure (3.11).

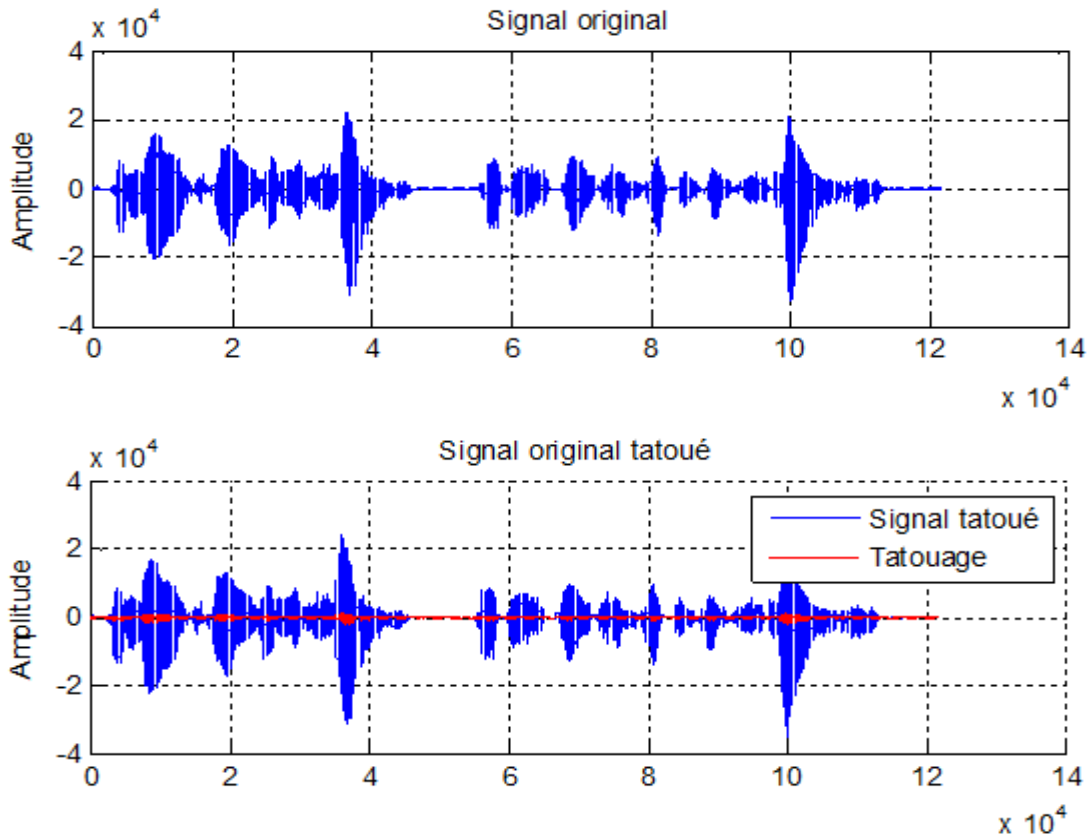


Figure3.11: tatouage global du signal audio.

III.15 Extraction et vérification du tatouage:

Garantir la vérification du tatouage est sans doute la contrainte première d'un système de tatouage. Le signal audio tatoué doit être identique au signal audio original d'un point de vue perceptif. Seul un ensemble de tests d'écoute réalisés par un grand nombre d'auditeurs peut établir de manière fiable la transparence du tatouage.

Pour l'extraction, le signal audio original n'est pas nécessaire. Avant de commencer la détection, le signal est filtré par un filtre passe bas. Le signal obtenu x_2 est démodulé en le multipliant par le bruit utilisé lors de l'insertion de la signature. On calcule ensuite la corrélation entre la signature et le signal obtenu. Et la détection de la signature est exécutée après le décodage entropique avec les mêmes procédures d'insertion sont employées pour extraire la signature.

III.16 Signature représentée par un ‘Bruit Unique’:

La signature de tatouage audio peut être un bruit unique qui ajoutera avant l'émission du signal, la détection de bruit est exécutée après le décodage en niveaux de réception. Afin de ne pas augmenter le débit du signal, seuls les blocs marqués dont la taille est inférieure ou égale au bloc original sont sélectionnés.

III.17 Bruit unique généré aléatoirement:

Dans ce cas du tatouage: la marque est comme un bruit produit aléatoirement avec le signal audio, ce bruit peut commencer ou débute de ce signal, ou moyenne ou à la fin. Avec un spectre d'énergie limité par le seuil d'audition utilisé. Un bruit unique peut être généré aléatoirement grâce à la fonction '*rand*' (existe sous Matlab) ou pseudo-aléatoirement par l'algorithme de génération du bruit blanc.

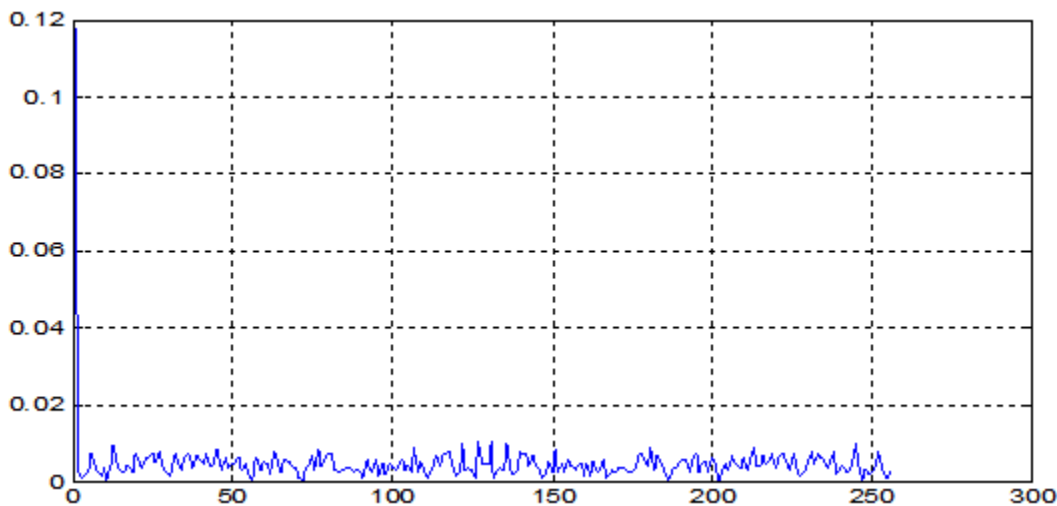


Figure 3.12: spectre de bruit.

III.18 Bruit unique généré pseudo-aléatoirement:

Le tatouage audio consiste à insérer une information inaudible dans un signal audio. Cette information est généralement représentée par un signal pseudo-aléatoire, le tatouage, détecté à l'aide d'une mesure de corrélation. Si le tatouage doit être robuste à des attaques malveillantes, le signal pseudo-aléatoire est impérativement secret et la détection est dite privée. Soit le vecteur x représentant le signal hôte original et le vecteur y représentant le signal tatoué après sa distorsion. Une clé secrète K est utilisée par une séquence pseudo-aléatoire u de moyenne nulle, elle est ajoutée au signal audio original. Le signal tatoué y est noté par :

$$y = x + mu \quad (3.8)$$

Avec N est la taille des vecteurs x et u . Le bruit n est considéré comme étant additif et le signal tatoué résultant s'écrit sous la forme suivante:

$$\hat{y} = y + n \quad (3.9)$$

La corrélation normalisée est calculée entre le signal tatoué et la séquence pseudo-aléatoire u pour extraire le tatouage.

III.19 Signature représentée par une séquence audio:

Le tatouage ou la signature de l'auteur qui est ajouté au signal audio peut être une séquence audio de ce signal tatoué, ou bien d'un autre signal audio.

Conclusion:

Le tatouage peut être utilisé donc pour améliorer les systèmes de transmissions existants. En effet, il peut être utilisé pour:

- Aider à récupérer le signal hôte transmis sur le canal de communication par la correction des erreurs, l'information cachée est utilisée pour réduire le bruit lié à la transmission.
- Tester la qualité de services des transmissions multimédia
- Evaluer, de manière subjective, la qualité du signal de parole transmis sur la bande téléphonique.
- Elargir la bande passante du signal transmis dans un système de communication.

Chapitre IV: Résultats et Evaluations

IV.1 Introduction :

Un système de tatouage audio doit répondre à un certain nombre de conditions pour qu'il soit considéré comme fiable et performant. Dans cette partie, nous complétons les résultats du chapitre précédent, tout en donnant les évaluations de notre procédure de tatouage, pour différents types de signaux.

IV.2 Fidélité perceptuelle:

La première condition pour notre système de tatouage, est qu'en aucun cas, le tatouage ne doit influencer sur l'information contenue dans le signal, ou sur l'écoute de ce dernier. C'est pour cela que nous avons analysé l'adaptabilité du tatouage aux différents types de signaux susceptibles d'être tatoués. En effet, un signal audio peut contenir des informations très différentes les unes des autres, nous avons donc étudié l'effet du tatouage dans différentes parties du signal audio :

IV.2.1 Partie voisée:

Comme dans la majorité des cas, la signature est représentée par un bruit, le masquage de ce dernier devient délicat quand il se trouve dans une zone voisée du signal. En effet, un bruit est plus difficilement masqué par un son voisé, qu'un son voisé par un bruit. Rappelons à titre d'illustration que l'indice de masquage est de l'ordre de 4 dB lorsque le son masquant est un bruit à bande étroite et le son masqué une sinusoïde, alors qu'il est d'environ 24 dB pour la configuration inverse [18].

Cela explique la discrimination des sons tonals et non-tonals dans le modèle psycho-acoustique de la norme MPEG1 'niveau1'. En effet, la détermination des seuils de masquage diffère selon la nature du son masquant ; les formules déterminées empiriquement pour le calcul de la fonction d'étalement sont différentes selon que la composante soit tonale ou non.

La figure suivante montre une parcelle d'un tatouage représenté par un bruit unique, masquée par une partie voisée du signal tatoué :

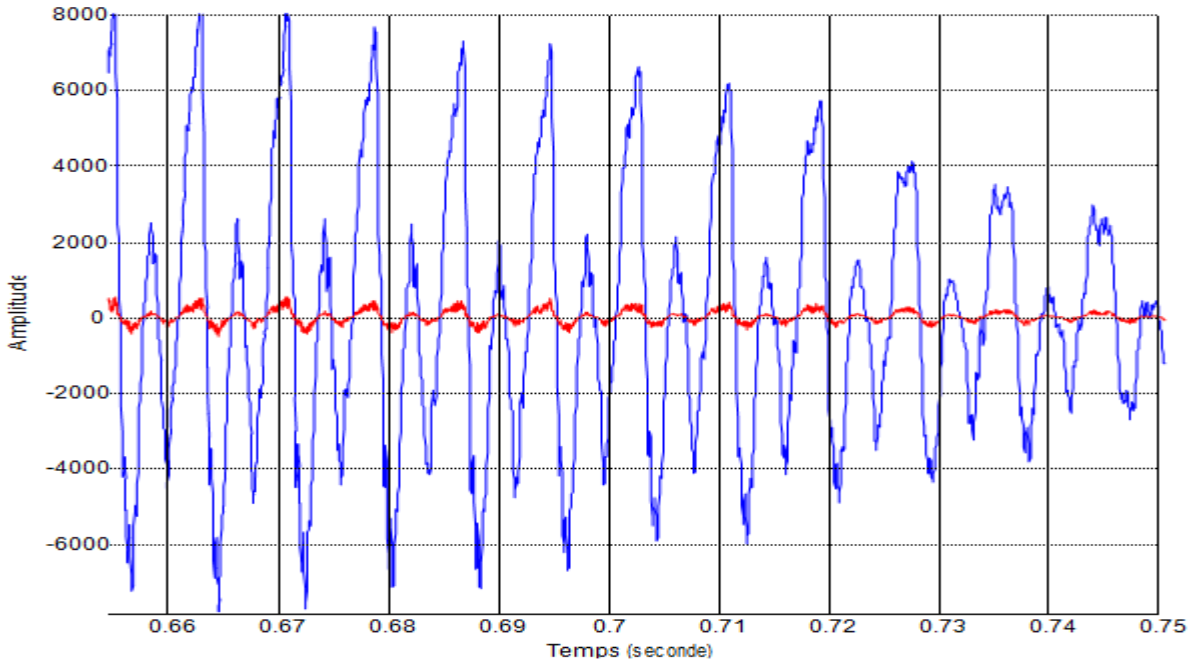


Figure4.1: Partie voisée du signal tatoué.

Nous observons sur cette figure que le tatouage en rouge évolue en fonction du signal, on remarque aussi une certaine périodicité dans le tatouage qui est en harmonie avec la périodicité du signal.

IV.2.2 Partie contenant du bruit:

Il est plus aisé à un bruit de masquer une signature, mais en général, dans le cas des séquences audio destinées au tatouage, les sons non voisés (sons contenant du bruit) sont d'amplitude faible, cela implique une forte atténuation de l'amplitude du tatouage. Pour une bonne résistance aux perturbations, le tatouage doit de préférence avoir les amplitudes les plus élevées possibles, tout en tenant compte de la condition d'inaudibilité. Ce compromis est assez difficile à réaliser.

La figure suivant illustre une portion d'un tatouage masquée par une partie bruitée (son non voisé) du signal tatoué :

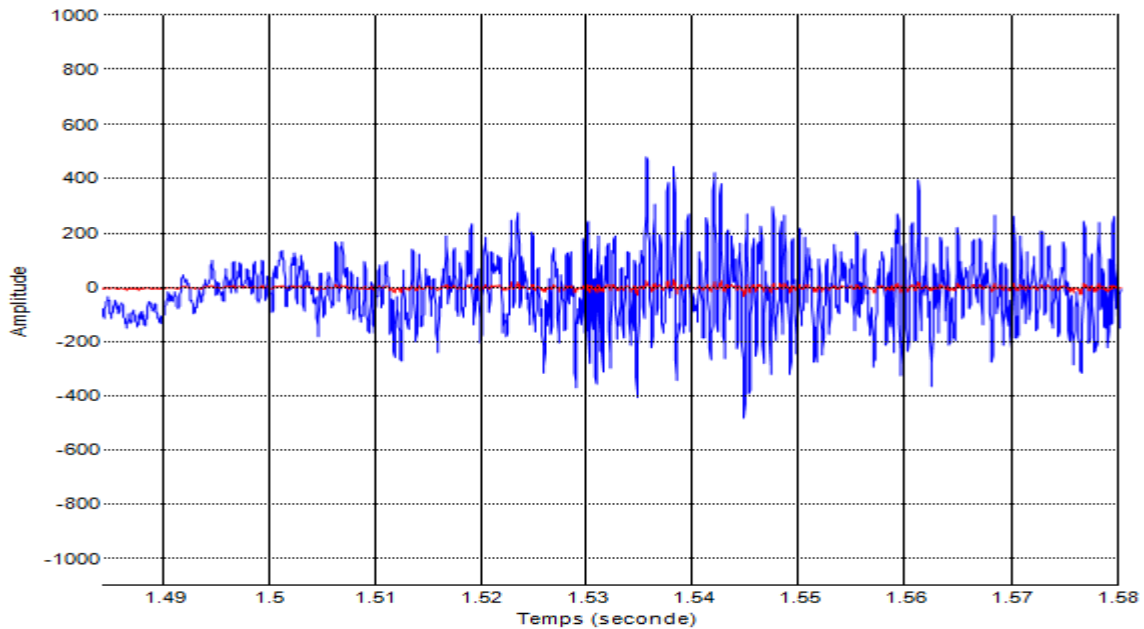


Figure4.2 : Partie bruitée du signal tatoué.

Nous observons que l'amplitude du tatouage est plus faible (indice ~ 40) comparativement au cas précédent (indice ~ 500), et que sa fréquence est plus élevée, suivant ainsi l'évolution des fréquences du signal.

IV.2.3Partie silencieuse:

Il est évident que sur une partie de silence d'un signal audio le tatouage ne peut être masqué, car dans de une telle zone, l'amplitude du tatouage est quasiment nulle, comme illustré sur la figure suivante :

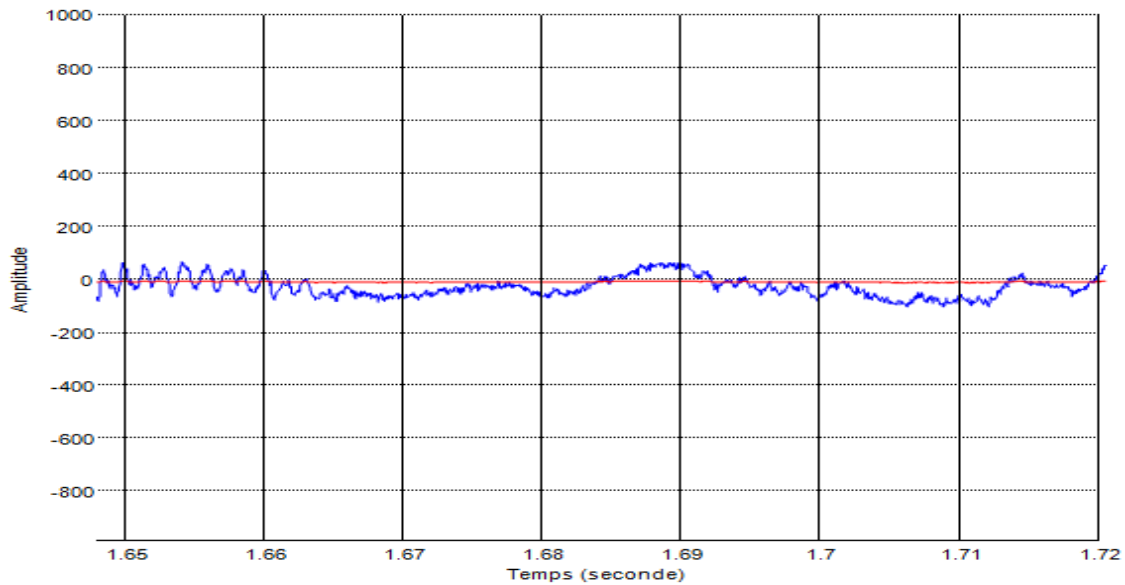


Figure4.3 : Partie silencieuse du signal tatoué.

Le tatouage sur cette figure a une amplitude quasiment nulle sur toute la parcelle silencieuse du signal.

A l'issue de ces observations, nous pouvons dire que notre tatouage s'adapte bien aux différentes natures de signaux que peut comporter une séquence audio ; cela lui permet d'être imperceptible dans tout les cas de figure. Cette propriété est due au double masquage fréquentiel – temporel.

En plus des observations précédentes, des tests d'écoutes sur différentes séquences (voix masculine, voix féminine, music...) de tailles différentes, nous permettent d'affirmer que l'inaudibilité du tatouage est assurée par notre algorithme.

IV.3 Fiabilité des résultats:

Pour prouver la propriété intellectuelle d'un média, la méthode utilisée doit être crédible et fiable. Le test de similitude de notre algorithme de simulation peut être considéré comme un outil crédible de jugement objectif.

IV.3.1 Cas sans transmission (local):

Après plusieurs tests, dans le cas idéal, c'est-à-dire qu'aucune perturbation n'est subie par le signal tatoué (sans transmission), notre test de similitude retourne le résultat :

$Sim(ew, w)=1$, c'est-à-dire 100% de similitude entre le tatouage extrait et le tatouage d'origine.

IV.3.2 Cas d'une transmission:

Dans le cas d'une transmission, des perturbations sur le signal tatoué doivent être prises en compte. Les résultats de similitude peuvent différer selon la chaîne d'opération que subit le signal (Compression, filtrage, codage, transmission...). Un seuil de tolérance : $S = \text{Sim}(ew, w)$ doit être fixé selon le cas de transmission et selon les besoins, en veillant à ce que le seuil doit être le plus proche possible de ce point n'a pas été testé en détail dans notre travail, nous le suggérons en perspective de travaux à venir.

IV.4 Robustesse:

Le tatouage doit être robuste aux attaques licites (compression, transmission...), comme aux attaques illicites (piraterie).

Dans le cas de la signature représentée par un bruit, le tatouage est statistiquement indétectable sans connaissances a priori. Concernant la signature représentée par une séquence audio propre à l'auteur, la détection de celle-ci par un pirate n'est pas impossible, mais reste difficile. De plus, le pirate n'aura atteint son but qu'en modifiant ou supprimant le tatouage, ce qui risque d'altérer le signal hôte et le rendre inexploitable.

IV.5 Coût et débit:

Le coût en termes de temps de calcul peut être un critère important, en particulier pour certaines applications destinées à fonctionner en temps-réel.

Après avoir effectué des tests de temps de calcul de notre algorithme, sur un système de puissance moyenne (CPU Dual core 2.5Ghz), nous avons obtenu les résultats suivants :

Tableau4.1 : Résultats des tests de temps de calcul sure différents signaux audio échantillonnés à 32Khz.

Duré de la séquence audio	Temps de calcul
3.8 secs	11 mn et 18 secs
3 secs	6 mn et 33 secs
2.2 secs	3 mn et 13 secs
2 secs	2 mn et 41 secs
1 secs	1 mn et 5secs

Nous observons que le temps de calcul n'évolue pas linéairement en fonction de la durée de la séquence audio à tatouer. De ce fait nous déduisons que le temps de calcul dépend aussi d'un autre paramètre, qui est la nature des sons tatoués (vu précédemment).

La moyenne du temps de calcul obtenue sur ce système est approximée à 2 minutes de calcul pour chaque seconde de séquence audio.

Concernant le débit ; notre but est de marquer le signal avec un tatouage dont l'amplitude est très faible, par une addition échantillon par échantillon. De ce fait, l'amplitude des échantillons augmente faiblement, et leur nombre ne changeant pas, le débit n'est donc pas affecté.

Conclusion:

Selon les résultats des différents tests que nous avons effectués, nous pouvons dire que notre schéma de tatouage répond aux exigences principales. Cependant il est nécessaire de procéder à des tests plus poussés, afin de déterminer ses limites, et d'effectuer d'éventuelles adaptations.

Conclusion générale

Ce mémoire traite du tatouage des signaux audio numériques, dans le contexte de la protection des droits d'auteur. Cette classe applicative permet l'identification de l'auteur ou du propriétaire légitime d'une œuvre audio numérique.

La finalité première de ce travail s'attache à la conception d'un tel système de tatouage. Cette conception est régie par la contrainte principale de l'inaudibilité, qui est contournée grâce au double masquage fréquentiel-temporel opéré par notre procédure de tatouage.

L'introduction dans ce travail d'un nouveau procédé de représentation de l'auteur, qui consiste à insérer une séquence audio à la place de la signature par un bruit, classiquement utilisée, permettra d'avoir un jugement subjectif à l'écoute qui est une alternative au jugement objectif habituellement exploité pour prouver la propriété. Ce nouveau procédé a été implémenté dans notre cas en Matlab et évalué sur quelques sons. Nous avons noté l'inaudibilité de la séquence audio insérée.

Comme perspective, nous suggérons des tests et évaluations plus élaborés sur différents types de signaux audio. L'étude précise du comportement de notre solution de tatouage, dans le cas d'une transmission bruitée, est aussi recommandée. Enfin, des tests d'écoute sur un échantillon de personnes (homme, femme) seront bénéfiques.

Bibliographie

- [1] CLEO BARAS: "Tatouage informé de signaux audio numériques", thèse de doctorat Ecole Nationale Supérieure des Télécommunications, (paris), 2005
- [2] LEANDRO de C. T. Gomes : Tatouage de signaux audio. Thèse de doctorat, Université Paris V, Paris, France, juillet 2002.
- [3] DALIL BATTIKH: "Sécurité de l'information par stéganographie basée sur les séquences chaotiques". INSA de Rennes, 2015.
- [4] SOFIANE BRACI: "Etude des méthodes de dissimulation informées de données appliquées aux supports multimédias". Thèse de doctorat, Université Paris Sud - Paris XI, 2010.
- [5] IMEN GHARBI: "Tatouage pour le renforcement de la qualité audio des systèmes de communication bas débit". Université René Descartes - Paris V; Ecole nationale d'ingénieurs de Tunis, 2013.
- [6] WAHID ABRKOUTI et ADNEN CHERIF: "Evaluation de Performances d'un Système de tatouage audio", 5th international conférence: Sciences of Electronique, Technologies of information and Télécommunications, TUNISIA, 2009
- [7] LEANDRO de C.T, et al: "Tatouage audio exploitant des propriétés de cyclostationnarité". Traitement du Signal 2001.
- [8] ABDRAHMANE, ADJILA: "Signature numérique pour fichiers audio", Mémoire de Magister, spécialité informatique, Université Kasdi Marbah à Ouargla. (Janvier 2013)
- [9] E.Zwiker, R.Feldtkelle: "Psycho-acoustique L'oreille récepteur d'information", Stuttgart, West Germany: HirZel Verlag, 1967 (Masson 19981, traduit de l'allemand par C.SORIN).
- [10] ASMAA AMEHRAÏ: "Dé-bruitage perceptuel de la parole", thèse de doctorat Ecole Nationale Supérieure des télécommunications de Bretagne. L'université Mohamed-V-AGDAL de rabat. (15 mai 2009)
- [11] B, T, ABDLLATIF: "traitement du signal audio dans le domaine codé: techniques et applications", thèse de doctorat Ecole National Supérieur de télécommunication (Paris), (2001)
- [12] GOYE, ALAIN: la perception auditive (cours P.A.M.U).Ecole Nationale Supérieure des télécommunications, (p27). (janvier 2002)

- [13] PH.GASSER: "Les formats MPEG audio", MSH Paris Nord-plats-forme Arts, sciences, Technologie. (Décembre 2006)
- [14] BUNITI." Traitement automatique de la parole en milieu bruite" : étude de modèles connexionnistes statiques et dynamiques. Université Henri Poincaré - Nancy 1,1997.
- [15] DOMINIQUE, GIBERT : "Eléments de Traitement du signal", (p7, p134), Août 1994
- [16] Alejandro Lobo Guerrero, Joël Liénard: "Implémentation d'un système de tatouage pour la transmission de données". [Consulté mars 2009]. Disponible sur <[http: www.lis.inpg.fr](http://www.lis.inpg.fr)>.
- [17] ITAN: "La compression des signaux audio et vidéo".2010.
- [18] T. Painter, A. Spanias: "Perceptual Coding of Digital Audio", Travaux de l'IEEE, Vol.88.No.4, 2000.

Résumé

L'avènement des documents sous format numérique offre des possibilités innovantes dans le domaine d'embarcation des données ; un document hôte, dans notre cas un signal audio, peut cacher une information numérique. La technique d'embarcation ou d'enfouissement est alors appelée tatouage.

Notre travail consiste à mettre en place une procédure de tatouage permettant d'intégrer une protection de droit d'auteur pour des données audio numériques, par modification directe des échantillons audio. Cette procédure exploite directement le masquage perceptuel, temporel et fréquentiel, pour garantir que le filigrane numérique (tatouage) soit inaudible et robuste.

Les Mots clés : Tatouage audio, Masquage perceptuel, psycho-acoustique

Abstract

The digital is document done to offer the innovative possibilities in the professional field of data, added document or in a case of an audio signal, can hide the digital data using the technique called (watermarking).

Our work is aimed at setting up a place copyright protection of audio digital data through direct modifying in the audio sample. This procedure is done directly by making a temporal, frequency, perceptual masking in order to ensure that watermarking is inaudible and robust.

Keywords: Watermarking, Masking perceptual, psychoacoustique

ملخص

تأتي الوثيقة الرقمية لعرض امكانيات مبتكرة في مجال الحرفي للمعلومات, وثيقة مضافة او في حالة إشارة صوتية, يمكن إخفاء المعلومة الرقمية بتقنية الإخفاء التي تسمى الوشم.

يتألف هذا العمل في وضع مكان لإجراء الوشم الذي يسمح بحماية حقوق المؤلف للأصوات الرقمية, من طرف التغير المباشر في العينة الصوتية و يتم هذا الإجراء مباشرة بالإخفاء السمعي ويكون زمني, ترددي من أجل ضمان وشم غير مسموع وقوي.

الكلمات المفتاحية : الوشم الصوتي , الإخفاء السمعي, علم السمعيات