

SYSTEME DE TATOUAGE AUDIO INFORME ET ITERATIF

A. LOBOGUERRERO , J. LIENARD , P. BAS

Laboratoire d'Images et Signaux

961 rue de la Houille Blanche, Domaine Universitaire BP 46

38042 Saint Martin d'Hères Cedex FRANCE

alobogue@lis.inpg.fr, joel.lienard@lis.inpg.fr, patrick.bas@lis.inpg.fr

Résumé –Le tatouage d'audio est une méthode qui permet l'insertion d'un message imperceptible dans un flux audio. Bien que le tatouage soit souvent utilisé pour garantir les droits d'auteur, il peut être aussi utilisé pour augmenter l'information transmise dans un contexte de communications. Dans les applications typiques de tatouage, le signal audio qui sert de support est considéré comme du « bruit », néanmoins, ce signal n'est pas aléatoire mais complètement connu lors de la transmission. L'idée principale de ce papier est d'utiliser cette propriété pour à la fois réduire l'imperceptibilité et augmenter la robustesse du système de tatouage. A partir de cette idée, nous proposons un model à partir d'une technique classique de modulation puis nous profitons de la connaissance du signal support pour réaliser la conception un système de tatouage « informé ». Nous avons modifié notre système à l'aide d'un contrôle itératif de puissance et d'un filtrage optimal. Nos avons testé ce système de tatouage informé face à différents traitements et nous l'avons comparé aux systèmes non informés fondés sur le filtrage adapté.

Abstract – *Audio watermarking is a method that allows the insertion of an imperceptible mark on an audio data content. Although the watermarking is often used to guarantee copyrights, it can also be used to increase the information transmitted in a communication context. In typical watermarking applications, the audio data content represents the host that supports the embedded information and it is considered as "noise". However, we can use the fact that this signal is not random but completely known for the transmitter to reduce imperceptibility and improve robustness. Based on this idea, we derive a model from a classical data transmission technique and we take advantage of the knowledge of the support information to design an informed embedding scheme. This model has been then modified by iterative controlling of the transmitted power and by the use of an optimal filter. This informed embedding scheme allows us to provide a robust data-hiding system to several treatments, and it's compared to non informed and matched filter embedding schemes.*

1. Introduction

Le tatouage de signaux [1] a été dans un premier temps présenté comme une technique permettant de répondre au problème de la protection de droits d'auteur. L'idée fondamentale consiste à ajouter une marque (également appelée une signature) dans les données à protéger. Cette marque doit rester inaudible : aucune distorsion n'est permise sur les signaux traités. Dans un tel contexte, le tatouage audio doit également être robuste au traitement habituel des signaux aussi bien qu'aux attaques malveillantes. En plus de son utilisation pour la protection de droits d'auteur, les techniques de tatouage peuvent être aussi utilisées pour la transmission de données [2]. On parle dans ce cas de "contenus augmentés". Plusieurs applications sont possibles dans ce contexte. Par exemple, si comme dans le projet ARTUS [3], l'information prioritaire à transmettre est la voix ou la parole, alors il peut être utile de transmettre l'information visuelle d'analyse telle que des lèvres déplaçant les paramètres, le visage du parleur ou n'importe quelle autre information identifiant le visage parlant. Ce projet vise à inclure des paramètres d'animation à la fois dans les contenus audio et visuel. Dans article, nous présentons une technique de

tatouage audio en donnant une particulière attention au taux de transmission du message inséré. Dans les systèmes de tatouage additifs, à différence des systèmes de communications classiques, la détection est en effet perturbée par le signal qui sert de support à l'information. L'idée principale de ce travail est d'utiliser le signal de support (disponible à la transmission) pour optimiser les conditions du système. Cela correspond à ce que plusieurs auteurs appellent le tatouage informé [4].

L'article [5] utilise cette idée. Nous présentons une méthode plus élaborée et montrons son efficacité par des tests.

2. Implémentation

2.1 Schéma Général

La méthode de tatouage proposée est fondée sur la technique de télécommunications CDMA (Access Multiple par Division de Codes)[6]. Cette technique permet des communications Multi-Utilisateurs simultanées et peut être également employée pour les applications mono-utilisateurs à faible RSB (Rapport Signal à Bruit). Cette méthode est basée

sur la modulation par étalement de spectre. Cette modulation permet une distribution d'énergie homogène dans toute la largeur de bande disponible. A partir de cette technique nous allons présenter l'insertion du message et le type de détection utilisé.

2.1.1 Insertion

La technique de modulation choisie utilise un seul code pour insérer un message binaire. La version positive du code est utilisée pour convoyer le bit 1 et, la version négative sera employée pour représenter le bit 0. La figure suivante illustre cette modulation.

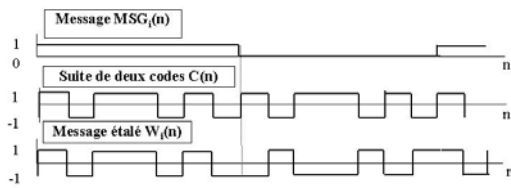


FIG.1 : Modulation par étalement de spectre à un seul code

Le code étalé peut être exprimé comme le produit

$$W(t) = MSG(t) \times C(t) \quad (1)$$

Dans (1), $MSG(t)$ représente le message à envoyer, $C(t)$ le code et $W(t)$ est le signal modulé obtenu par ce processus. Le signal $W(t)$ est additionné au signal sonore pour générer le signal tatoué $S_w(t)$.

$$S_w(t) = S(t) + W(t) \quad (2)$$

Afin de diminuer l'audibilité du message, les codes insérés doivent également satisfaire plusieurs conditions en ce qui concerne la puissance et la distribution spectrale. Ces conditions seront étudiées dans cet article.

2.1.2 La Détection Optimale.

A l'étape de réception, le RSB a été maximisé. Ceci est fait en complément de techniques classiques de tatouage par étalement de spectre dans lesquelles le bruit est considéré comme étant blanc et gaussien. Dans l'application en étude, le signal acoustique joue le rôle d'un bruit. En fait, ce « bruit » (musique, parole) a un comportement très dynamique de fréquence et il sera traité d'une manière optimisée à l'aide du Filtrage Optimal [7].

Ce filtre maximise le SNR dans le cas de signaux immergés dans un bruit non blanc. L'équation qui caractérise ce filtre est:

$$Fop(f) = \frac{C^*(f)e^{2\pi fT}}{S_s(f)} \quad (3)$$

Ce type de filtre peut être décrit en deux parties. La première a été notée comme $C^*(f)e^{2\pi fT}$ et correspond au filtre

adapté au signal $C(t)$. L'argument de l'exponentielle représente un retard temporel et T est la durée du signal. La deuxième partie, noté $S_s(f)$ est la Densité Spectrale de Puissance du signal qui perturbe la détection. Ce filtre correspond à un filtre adapté pour lequel le bruit est blanchi.

La détection est faite après la convolution du signal reçu (signal envoyé + bruit du canal) avec le Filtre Optimal. Cette convolution est échantillonnée toutes les T secondes de manière synchronisée. La décision sera prise par comparaison de cette valeur avec un seuil prédéterminé. Le résultat de cette opération sera nommé $Y(t)$.

Dans ce système la valeur du seuil est zéro (message binaire). La loi de décision est décrite par les équations suivantes :

$$\begin{aligned} \text{If } Fop(f) \otimes S_w(f) \Big|_{t=T} > 0 &\Rightarrow \text{bit détecté} = 1 \\ \text{If } Fop(f) \otimes S_w(f) \Big|_{t=T} < 0 &\Rightarrow \text{bit détecté} = 0 \end{aligned} \quad (4)$$

2.2 Améliorations Proposées

En s'appuyant sur le schéma général, nous proposons plusieurs traitements permettant d'optimiser le système. L'idée principale de cette implémentation s'appuie sur le fait que le signal audio est disponible lors de la transmission. Nous avons déjà souligné que dans les systèmes de tatouage cette information est traitée comme du bruit. Néanmoins, cette information est utilisable d'une manière déterministe dans l'étape d'insertion pour augmenter les performances de détection. Cette opération s'effectue par un système de tatouage informé. Dans la suite nous expliquerons les quatre étapes du système et la nécessité d'utiliser un système itératif.

2.2.1 Contrôle de la puissance des codes

Nous cherchons à déterminer le débit maximum qui permettra une détection sans erreur et une gêne auditive la plus faible possible. En ce qui concerne l'audibilité, nous devons limiter l'effet de l'ajout du message. Pour cela, il semble pertinent de minimiser la puissance du message ajouté. L'idée est de calculer à partir de l'information disponible en transmission (le support audio) la puissance minimale à transmettre pour obtenir une détection sans erreurs. Le calcul doit être réalisé en concordance avec les équations présentées en (4) Ceci doit être fait pour chacun des débits étudiés. Ce contrôle de puissance est possible grâce à la connaissance du signal audio dans l'étape de transmission et nous permet de réaliser une approximation déterministe du signal de support et de calculer les valeurs désirées. L'utilisation de cette puissance en complément d'un modèle psycho-acoustique garantit une détection sans erreurs ainsi que une perception minimale du message transmis. Afin de trouver cette puissance, il est nécessaire de déterminer la valeur minimale du coefficient par laquelle chacun des codes

sera multiplié. A partir des équations présentées en (4), le calcul rend le résultat suivant :

$$K > - \frac{\int_{\tau} S(\tau) Fop(t-\tau) \partial \tau}{\int_{\tau} W_m(\tau) Fop(t-\tau) \partial \tau} \Bigg|_{t=T} \quad (5)$$

Cette valeur est calculée pour le cas d'un canal idéal sans bruit additif. Pour prendre en compte d'une éventuelle présence du bruit, nous avons augmenté légèrement cette valeur par un montant $\Delta \epsilon$.

Dans cette équation C_m dénote le code pondéré par le model psycho acoustique, F_{op} est le même qui a été présenté en (3). La figure No. 2 illustre ce concept dans la région nommée « Tatouage Informé »

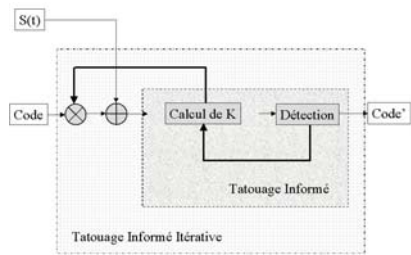


Figure No.2 Tatouage Informé Itératif

2.2.2 Ajustement itératif du contrôle de puissance

Nous avons présenté (cf. section 2.1.2.) un système de détection fondé sur la technique de Filtrage Optimal. Ce filtre comporte une étape de blanchiment du signal support suivi du filtre adapté au code utilisé pour réaliser la modulation. Pour réaliser le blanchiment, nous avons besoin de connaître la densité spectrale de puissance du signal de support audio. Cette information n'est plus disponible car nous avons modifié le signal support lors de l'ajout du message. Bien que différence apportée par cette modification soit faible, elle peut entraîner un nombre important d'erreurs.

L'idée est donc calculer à nouveau la puissance nécessaire à partir de la Densité Spectral de Puissance disponible. Pour cela, nous avons modifié le calcul de K à partir de la différence existante entre la DSP disponible et désirée. L'idée est de recalculer la puissance nécessaire à partir la DSP modifiée (obtenue de $S_s+W(f)$). Ce développement nous amène à l'implémentation du système itératif montré dans la figure 2 (région Tatouage Informé Itératif). Le module « calcul de K » montré dans cette figure est celui qui a été décrite par (5). En calculant chaque itération, le système affecte le calcul jusqu'à l'obtention d'une détection sans erreurs.

Les équations résultantes sont:

$$K_{i+1} > - \frac{\int_{\tau} S(\tau) Fop_i(t-\tau) \partial \tau}{\int_{\tau} K_i W(\tau) Fop_i(t-\tau) \partial \tau} \Bigg|_{t=T} \quad (6)$$

$$Fop_i(t) = TF^{-1} \left[\frac{C^*(f) e^{2\pi j f t}}{S_{s+K_i W}(f)} \right] \quad (7)$$

2.2.3 Pondération Spectrale des codes à l'aide du modèle psycho acoustique

En complément du *contrôle de puissance des codes* (sous-section 2.2.1) nous avons décidé de réaliser l'implémentation d'une autre stratégie afin de diminuer l'audibilité du message ajouté. Pour faire ceci, il est nécessaire d'assigner les coefficients spectraux de codes dans les fréquences où l'oreille humaine est la moins sensible. Nous avons donc utilisé réalisé une pondération spectrale à l'aide d'un modèle psychoacoustique dérivé de la compression Mpeg-1 standard.

2.2.4 Etape de synchronisation

Pour réaliser la synchronisation un code supplémentaire est ajouté. La technique classique de corrélation glissante est utilisée pour déterminer l'instant d'apparition de ce code et ainsi permettre une synchronisation.

3. Implémentation définitive

Nous avons expliqué toutes les étapes associées au processus de tatouage. Nous pouvons présenter le schéma définitif dans la figure No.3

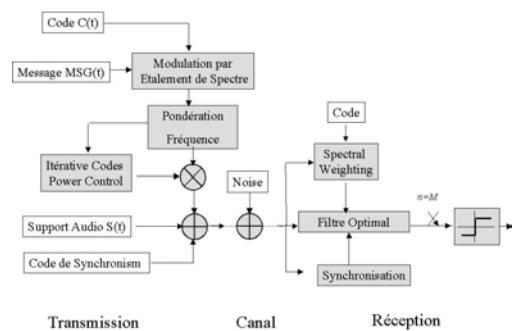


Figure No.3 Implémentation définitive

4. Résultats

Pour évaluer les performances du système, nous avons testé le système en deux étapes principales. Dans la première étape nous avons déterminé le débit maximum pour lequel le message reste inaudible. Pour cela, nous avons pris un échantillon de dix secondes de musique et nous avons ajouté le message. Nous avons ensuite réalisé une évaluation simple

de l'audibilité du message inséré pour différents débits. Pour un débit de 86 bit/s le message la gêne introduite par l'ajout du message est suffisamment faible pour être acceptable. Pour des débits supérieurs cette gêne est de plus en plus perceptible.

Dans la deuxième partie, nous avons fixé ce débit et nous avons décidé de comparer le schéma proposé avec les techniques de tatouage traditionnelles. Nous avons décidé d'utiliser comme critère de comparaison le RSB à la place de l'audibilité du message introduit. Le calcul des rapports signal à bruit a été donc réalisé dans plusieurs cas différents.

Nous avons réalisé un premier test afin de comparer le système de control de puissance informé avec l'implémentation classique (boucle ouverte à puissance constante). Nous avons utilisé la puissance constante nécessaire pour obtenir le même nombre d'erreurs et nous avons mesuré le rapport signal à message.

Dans le deuxième test, nous avons utilisé la méthode informée itérative présentée ci-dessus mais le filtre optimal a été remplacé par le système de filtrage adapté.

Le troisième test correspond à la procédure proposée complète (Tatouage Audio Informé et Itératif).

La comparaison entre le rapport de puissances du premier et troisième test montre le différences de performances entre système de contrôle de puissance informé avec l'implémentation classique (boucle ouverte à puissance constante).

Une comparaison de résultats du troisième test avec ceux du premier test permet d'évaluer les performances des filtres adapté et optimal.

Dix signaux ont été analysés et les résultats obtenus ont été évalués pour un débit de 86 bit/s. Les sons utilisés sont échantillonnés à 44.1kHz sur 16 bits. Chacun de fichiers résultants a une durée de 116 seconds. Ceci correspond à 10.000 bits d'information envoyés.

Le tableau No. 1 présente la synthèse des résultats. Pour la mesure du RSB nous avons pris la puissance du signal (support audio) divisé par la puissance du message ajouté en dB's.

avons simulé un canal gaussien additif avec un RSB de 42 dB's. Afin de considérer cette présence de bruit, nous avons modifié la valeur de $\Delta\epsilon$ présenté en 2.2.1. jusqu'à obtenir un obtenir un BER de 0.82. Dans ces conditions le message ajouté reste également inaudible.

5. Conclusion et Perspectives

Le système de tatouage audio fondé sur la connaissance déterministe du signal a été brièvement présenté. Nous avons vérifié que le filtrage optimal est plus performant que le simple filtre adapté en mesurant la puissance nécessaire pour obtenir une transmission sans erreurs. Nous avons souligné aussi que le signal de support peut être traité différemment du bruit. Le système de Tatouage Audio Informé et Itératif a été développé sur cette idée. Des débits d'insertion importants peuvent être obtenus grâce à cette technique. De manière analogue, l'idée d'utiliser l'information contenu dans le signal support sera utilisée prochainement pour aider à la synchronisation du système.

6. Références

- [1] V. Basia and I. Pitas. Robust Watermarking in time domain" Proc EUSIPCO, volume 1, Rhodes, 1998.
- [2] A. LoboGuerrero, P. Bas, J. Liénard, "Implémentation d'un système de tatouage pour la transmission de données" JEP 2002, pp. 93-96. Nancy, France.
- [3] Internet site: http://www.telecom.gouv.fr/rnrt/projets/res_01_37.htm
- [4] M.L. Miller, J. Cox, J.A. Bloom, "Informed Embedding: Exploiting Image and Detector Information During Watermark Insertion," IEEE International Conference on Image Processing, vol. 3, pp. 1-4, 2000.
- [5] Transmission d'un message numérique caché dans un signal audio. Joël Liénard. Colloque GretsI 2001, Toulouse, France.
- [6] J.G. Proakis, Digital Communications, McGraw-Hill, New York, 1995.
- [7] J. Max, Méthodes et Techniques de Traitement du Signal et Applications aux Mesures Physiques, Vol II, Masson, Paris, 1986

Schéma Proposé	Schéma Filtre Adapté	Puissance Constante
Filtrage Optimal	Filtre Adapté	Filtrage Optimal
Contrôle Itératif	Contrôle Itératif	Sans Contrôle Itératif
63,7	61,8	43,2

Tableau No. 1 RSB obtenus

Bien que cette implémentation garantisse une détection sans erreurs dans le cas de canal idéal sans bruit additif, nous