Tatouage informé hiérarchique d'un message hiérarchique (en vue de la protection vidéo)

Nicolas Tournier, Marc Chaumont, William Puech

Tatouage informé hiérarchique d'un message hiérarchique (en vue de la protection vidéo)

N.Tournier M.Chaumont W.Puech

LIRMM, Université de Montpellier II, UMR CNRS 5506
161, rue Ada, 34392 MONTPELLIER CEDEX 05, FRANCE

{Nicolas.Tournier, Marc.Chaumont, William.Puech}@lirmm.fr

Résumé

Dans cet article, nous proposons une méthode de tatouage informée hiérarchique sur des images fixes, pouvant s'adapter à la vidéo. A partir d'un message dont l'information est hiérarchisée, nous souhaitons la dissimuler de manière hiérarchique dans le document hôte. Ainsi l'acquéreur du document haute résolution aura accès à l'ensemble des données insérées, contrairement à celui qui possède une résolution inférieure où l'extraction ne sera que partielle.

Pour ce faire trois types de tatouage basés sur QIM[1] *[1] sont évalués sur des supports d'insertion différents (séparés, emboîtés et glissants). Les résultats obtenus montrent que le tatouage hiérarchique de message hiérarchisé peut être réalisé de manière imperceptible et robuste, et que les espaces emboîtés et glissants assurent une plus grande robustesse.*

Mots clefs

Tatouage informé, tatouage hiérarchique, suivi de copie, traçabilité

1 Introduction

Les documents vidéos hiérarchiques sont des supports numériques pouvant s'adapter en résolution spatiale, temporelle, en qualité et en débit. L'information hiérarchisée peut être distribuée sur un téléphone portable ou un écran haute définition par exemple. Dans le but de protéger ce type de document et d'enrichir son contenu, nous nous intéressons au tatouage informé hiérarchique de messages hiérarchiques.

Le tatouage hiérarchique (*scalable watermarking*) n'est pas encore clairement défini pour l'ensemble de la communauté scientifique. C'est en 1998, que Wang et Icuo [2] introduisent cette notion. Depuis 2003, la définition qui fait référence est celle de Piper *et al.* [3] mais ne prend en compte que des techniques d'insertion par étalement de spectre. Elles sont non-informées, donc de faible capacité et peu robustes.

1. QIM : Quantization Index Modulation

Dans nos travaux, nous considérons des documents hiérarchiques en résolution spatiale et des méthodes d'insertion informées. Par tatouage hiérarchique, nous entendons que la marque doit être détectable dans un ensemble fixé de résolutions.

Dans cet article, nous détaillons en section 2, la méthode développée et nous présentons, en section 3, les résultats obtenus de manière expérimentale. Enfin nous concluons en section 4.

2 Méthode Proposée

Dans cette section, nous présentons les objectifs et les stratégies de hiérarchisation de nos méthodes (2.1). Puis, nous rappellons brièvement l'algorithme QIM (2.2) et nous décrivons nos algorithmes d'insertion (2.3) et d'extraction (2.4) des messages hiérarchiques en utilisant nos supports d'insertion.

2.1 Objectifs et stratégies de hiérarchisation

Le but de la méthode est de proposer une technique de tatouage, hiérarchique et informée, d'un message dont l'information est hiérarchisée, sur des images. Supposons qu'un client ait fait l'acquisition d'un film tatoué parmi différentes résolutions : HD 1080i, HD 780p, DVD (DV-PAL), etc... Dans le cadre d'une application de traçabilité (*fingerprinting*), l'information primordiale qui nous intéresse concerne les coordonnées du client (nom, prénom, date d'achat), et ce qui caractérise le document acheté correspond à de l'information secondaire (titre du film, réalisateur, résumé).

Ainsi, nous pouvons découper le message en trois paquets. Par exemple, dans le premier nous considérons les informations de l'utilisateur, puis dans le second les principales métadonnées du film (titre, réalisateur) et dans le dernier un bref résumé.

En fonction de la résolution obtenue, l'acquéreur n'a accès qu'à une partie des informations. En gardant le contexte du suivi de copie, il est nécessaire de pouvoir détecter, quelle que soit la résolution, les coordonnées du client. Les informations supplémentaires sont considérées comme des métadonnées qui enrichissent le contenu du média. En acquérant un fichier de résolution inférieure une partie des

métadonnées n'est plus accessible mais les identifiants de l'acquéreur restent détectables.
Nous avons porté notre attention sur une approche QIM proposée par Chen et Wornell [1] car l'algorithme est rapide et peut s'adapter aisément à du tatouage de vidéos.
Nous supposerons, que nous avons une image hiérarchique décomposée en n niveaux de résolution. Nous souhaitons donc insérer un message hiérarchique et sur les résolutions inférieures pouvoir en extraire une partie. L'image est décomposée en n niveaux dans le domaine des ondelettes Daubechies 9/7 [4]. Ces niveaux sont caractérisés par les coefficients des sous-bandes $\{HH, LH, HL\}$ et sera noté $\forall i \in \{1, \ldots, n\}$, le i^{eme} niveau $X_i = \{HH_i, LH_i, HL_i\}$ (*cf.* Figure 1).

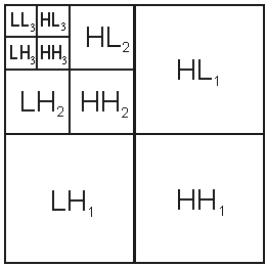


Figure 1 – *Schéma d'une décomposition en ondelettes sur trois niveaux.*

Concernant le message m, nous supposons qu'il est hiérarchique dans le sens où nous le découpons en k paquets $(m_1, \cdots, m_k)$ (avec $k < n$) qui seront dissimulés dans certains niveaux de la décomposition. Nous décomposons l'image en n niveaux et nous nous intéressons aux ensembles de coefficients X_i $(\forall i \in \{1, \cdots, n\})$. Dans ces conditions nous pouvons dissimuler le message m_i dans X_i. Pour notre approche, dans le domaine du tatouage hiérarchique, nous avons développé trois méthodes d'insertion. Nous définissons $Supp(i)$ comme le support d'insertion du message, connaissant les niveaux $\{X_{n-i+1}, \cdots, X_n\}$. En d'autres mots, c'est l'ensemble des coefficients qui seront utilisés pour la dissimulation des données.
Nous présentons les différentes constructions des supports d'insertion des paquets, pour commencer nous introduisons ce que nous appellerons des supports séparés, puis les supports emboîtés et pour finir les supports glissants.

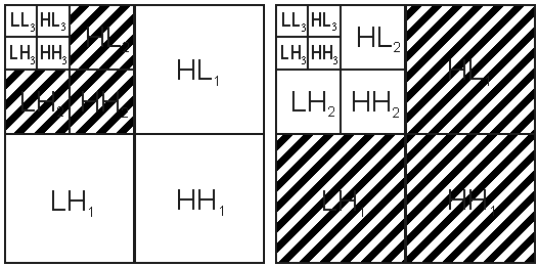
Méthode des supports séparés. Lorsque nous tatouons sur des niveaux séparées (*cf.* Figure 2), d'après les notations nous avons trivialement :

$$Supp(i) = X_{n-i+1}.$$

Méthode des supports emboîtés. De manière hiérarchique, nous pouvons choisir de tatouer sur l'ensemble des coefficients ondelettes mis à notre disposition connaissant les couches $\{(n-i+1), \cdots, n\}$, de manière à toujours insérer une partie de l'information dans les basses fréquences :

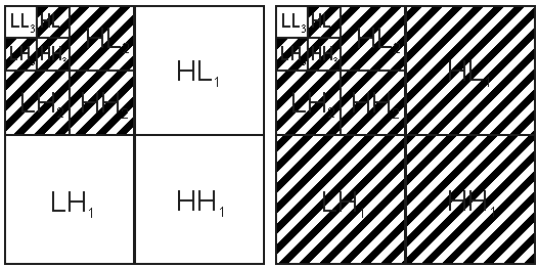
$$Supp(i) = \bigcup_{j=1}^{i} X_{n-j+1}.$$

Exemple : L'image est décomposée en trois niveaux ($n = 3$), $Supp(2) = X_3 \cup X_2$ et $Supp(3) = X_3 \cup X_2 \cup X_1$ (*cf.* Figure 3).



(a) Ex : $Supp(2)$ (b) Ex : $Supp(3)$

Figure 2 – *Représentation graphique des supports pour la méthode d'insertion sur des supports séparés.*



(a) Ex : $Supp(2)$ (b) Ex : $Supp(3)$

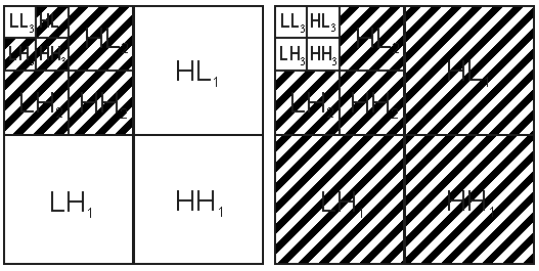
Figure 3 – *Représentation graphique des supports pour la méthode d'insertion sur des supports emboîtés.*

Nous notons que $Supp(1) \subset Supp(i)\ \forall i$, ce qui signifie que toute insertion dans le support aura une influence sur les sous-bandes des plus basses fréquences dans l'optique de conserver une certaine robustesse. De plus $Supp(i) \subset Supp(i+1)\ \forall i$, nous avons une suite d'ensembles emboîtés et croissante au sens de l'inclusion et par conséquent de dimension croissante, ce qui permet à l'information d'être diffusée au fur et à mesure sur l'ensemble des coefficients.

Méthode des supports glissants. Nous pouvons également conserver une intersection des espaces d'insertion en nous fixant un paramètre $p \geq 1$ comme étant le nombre maximal de support auquel appartient une sous-bande :

$$Supp(i) = \bigcup_{j=i}^{i+p} X_{n-j+1}.$$

Exemple : Nous fixons $n = 3$ et $p = 1$, $Supp(1) = X_3 \cup X_2$ et $Supp(2) = X_2 \cup X_1$ (*cf.* Figure 4)



(a) Ex : $Supp(1)$ $(p = 1)$ (b) Ex : $Supp(2)$ $(p = 1)$

Figure 4 – *Représentation graphique des supports pour la méthode d'insertion sur des supports glissants.*

Contrairement au cas précédent, nous n'avons plus de suite d'ensembles croissante au sens de l'inclusion. Cependant, pour les mêmes raisons qu'évoquées précédemment, nous avons gardé une suite de dimension croissante. En effet $|Supp(i+1)| - |Supp(i)| = |X_{i-p-i}| - |X_i + 1| > 0$. En vue d'un tatouage hiérarchique, nous avons posé ces ensembles pour avoir un compromis entre robustesse et imperceptibilité. L'approche est plus robuste qu'en tatouant séparément chaque bande et dégradant moins l'image par comparaison avec les supports emboîtés.

2.2 Algorithme QIM

Dans le cas du tatouage informé, nous avons besoin d'un dictionnaire Σ qui regroupe l'ensemble des mots de code. Ce dictionnaire est représenté par une latice de dimension N notée $(\Lambda, +)$. Une latice est un sous groupe de $(\mathbb{R}^N, +)$ qui peut se caractériser de manière unique par une matrice génératrice $G \in \mathcal{M}_N(\mathbb{R})$:

$$\Lambda = \{Gx : x \in \mathbb{Z}^N\}.$$

Nous définissons ensuite un ensemble de sous-latices Λ_i tel que $\forall i\ \Lambda_i \subset \Lambda$. Chaque Λ_i représente un message du dictionnaire $\Sigma = \{0; 1\}$ ce qui fait que nous avons une représentation d'un message par latice. L'espace $\mathbb{R}^N$ est alors "pavé" par la latice Λ. Ainsi nous avons une infinité de représentations du dictionnaire (et donc de notre message) dans $\mathbb{R}^N$. Cette infinité de représentations d'un même message dans un espace quelconque est caractéristique des techniques de tatouage informé :

$$\Lambda = \bigcup_i \Lambda_i.$$

Dans l'algorithme QIM [1], chaque message m_i est représenté par une sous-latice Λ_i. Le but de la méthode est de trouver la plus proche du signal hôte X codant le message m_j. Le vecteur hôte X est alors déplacé dans la sous-latice Λ_j (*cf.* Algorithme 1). Lors de la détection, il suffit de relever à quelle sous-latice appartient le signal reçu et ainsi de connaître le message transmis.

Algorithme 1 : Algorithme d'insertion QIM

Entrées : Une image $X \in \mathbb{R}^N$, un message $m_j \in \{0; 1\}^L$ et une clé secrète K

Résultat : L'image tatouée $Y \in \mathbb{R}^N$

début

Construire la latice Λ et les sous-latice Λ_i ;

Déplacer le vecteur X dans la sous-latice la plus proche Λ_j, nous obtenons le vecteur Y ;

fin

2.3 Algorithme d'insertion du message

Nous détaillons dans cette section les hypothèses que verifient nos données lors de nos expériences et l'ensemble des pré-traitements avant l'insertion du message (transformations ondelettes, séparation des coefficients ondelettes, construction des supports et des lattices) qui décrivent nos algorithmes d'insertion de données (*cf.* Figure 5).

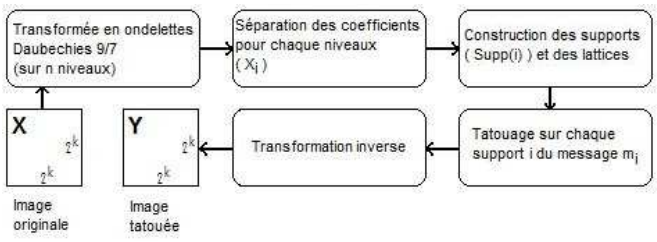


Figure 5 – *Schéma général de l'algorithme de tatouage.*

Données initiales. Nous disposons d'un message hiérarchique $m \in \{0; 1\}^L$ que nous découpons en k paquets $m = m_1, \cdots, m_k$ et d'une image en niveau de gris X que nous représentons sous forme matricielle, $X \in \mathcal{M}_{H,W}(\{0, 1, ..., 255\})$ avec H étant la hauteur de l'image et W sa largeur et $N = H \times W$ la dimension de l'image. La matrice X est la représentation de l'image dans le domaine spatial dont les coefficients sont les intensités lumineuses de chaque pixel. Nous supposerons que les images sont dyadiques.

Transformations. Avant de procéder à l'insertion, nous décomposons l'image I dans le domaine des ondelettes 9/7 Daubechies [4] en n niveaux (avec généralement $n \geq 2$). Ces décompositions hiérarchiques sont utilisées par l'algorithme de compression JPEG2000 [5]. Cette transformation est donc adaptée à notre problème de tatouage hiérarchique pouvant être à terme intégrée au sein d'un codeur vidéo.

Séparation. Une fois la transformation effectuée, nous séparons les coefficients en n sous-ensembles :

– LL_n : l'ensemble des coefficients contenant de l'information spatiale. Il s'agit des coefficients de l'image

d'approximation, une imagette qui est la représentation de l'image originale en très basse résolution ;
– X_i : l'ensemble des coefficients correspondant du i^{er} niveau de la décomposition ;

Cette séparation effectuée, nous obtenons n niveaux de décomposition, donc $(n+1)$ résolutions différentes. Nous reconstruisons la j^{eme} image par la transformée inverse en ondelettes à partir des coefficients $LL_n \cup \bigcup_{i=0}^{j-1} X_{n-i+1}$.

Construction des supports et des latices. La construction des latices est fortement liée à celle des supports présentés dans la section 2.1. En effet, la taille des porteuses dépend de la dimension de ce dernier. Par exemple, lorsque nous tatouons sur des supports séparés, la dimension des supports d'insertion est égale au nombre de coefficients du niveau $|X_i|$.
Sur chaque support, et quelle que soit la méthode utilisée, la construction des latices est analogue. Nous disposons d'un message hiérarchique $m \in \{0;1\}^L$, $m = m_1, \cdots, m_k$, à insérer dans une image $X \in \mathbb{R}^N$. Notre algorithme est basé sur l'algorithme QIM. Comme nous l'avons vu, nous construisons la latice Λ qui représentera notre dictionnaire Σ. Pour chaque espace d'insertion $Supp(i)$ nous générons pseudo-aléatoirement $|m_i|$ vecteurs normés ω_j de taille $|Supp(i)|$. D'après des résultats probabilistes, ces vecteurs sont deux-à-deux quasi-orthogonaux et vérifient l'équation suivante :

$$\forall x, y \in [1\ ;\ |m_i|]\ ;\ x \neq y\ ;\ (\omega_x|\omega_y) < ||\omega_x|| \cdot ||\omega_y||.$$

Algorithme. L'insertion du message se fait en appliquant l'algorithme QIM [1] en construisant pour chaque support d'insertion $Supp(i)$ une latice Λ où nous dissimulons une partie du message m_i. Il suffit ensuite d'effectuer les transformations inverses et retourner l'image tatouée. L'algorithme est plus détaillé dans l'Algorithme 2).

2.4 Algorithme d'extraction du message

La détection du message se fait de manière analogue à l'algorithme QIM. Nous nous ramenons à une image transformée par les même méthodes que lors de l'insertion, et par projection orthogonale, il suffit de déterminer dans quelle latice Λ_i se trouve la projection l'image reçue, nous retrouvons ainsi le message m_i correspondant. (*cf.* Algorithme 3).

3 Résultats expérimentaux

Pour chaque expérience effectuée, nous avons pris un jeu de 100 images de dimension 512×512 extraites de la base BOWS [2]. Pour chaque image, nous tatouons à un PSNR [3] de $45\ dB$ environ, et effectuons une décomposition en ondelettes sur deux niveaux ($n = 2$). Nous disposons d'un message hiérarchique m de 192 bits qui est inséré dans l'image (soit 96 bits par support).

2. BOWS : http ://bows2.gipsa-lab.inpg.fr/
3. PSNR : Peak Signal to Noise Ratio

Algorithme 2 : Algorithme d'insertion QIM hiérarchique

Entrées : Une image $X \in \mathbb{R}^N$, une clé secrète K et un message hiérarchisé $m \in \{0;1\}^L = (m_i)_{i=1\cdots k}$
Résultat : L'image tatouée $Y \in \mathbb{R}^N$

début

Transformer l'image X dans le domaine ondelettes Daubechies $9/7$ en n niveaux ;

Séparer les coefficients ondelettes en $n+1$ sous-ensembles LL_n, $X_i (i = 1 \cdots n)$;

Construire les espaces d'insertion $Supp(i)$;

pour *tout* $Supp(i)\ i = 1 \cdots k$ **faire**

Construire la latice Λ sur $Supp(i)$;

Insérer m_i en appliquant l'algorithme d'insertion QIM sur la latice Λ ;

fin

Effectuer les transformations inverses ;

Retourner Y l'image tatouée ;

fin

Algorithme 3 : Algorithme d'extraction QIM hiérarchique

Entrées : Une image $Y' \in \mathbb{R}^N$ et une clé secrète K
Résultat : Un message $m' \in \{0;1\}^L$

début

Transformer l'image Y' dans le domaine ondelette Daubechies $9/7$ en n niveaux ;

Séparer les coefficients ondelettes en $n+1$ sous-ensembles Y'_{SP}, $Y'_i (i = 1 \cdots n)$;

Construire les espaces d'insertion $Supp(i)$;

pour *tout* $Supp(i)\ i = 1 \cdots k$ **faire**

Construire la latice Λ sur $Supp(i)$;

Extraire m'_i en appliquant l'algorithme d'extraction QIM sur la latice Λ ;

Concaténer m'_i à m' ;

fin

Retourner m' ;

fin

Nous rappellons que notre message est hiérarchisé par ordre d'importance des informations. Pour des applications de traçabilité, les informations prioritaires sont les coordonnées du client. Elles sont dans un paquet qui est inséré dans le support le plus robuste. Si le paquet contient des informations moins importantes, alors il est tatoué sur un support moins robuste.
Avant l'extraction du message dans chaque niveau, nous étudions la robustesse de l'insertion face à trois types d'attaques : filtrage gaussien (3.1), attaque valumétrique (3.2) et compression JPEG (3.3). Ces attaques sont effectuées sur la haute résolution de l'image tatouée. Nous comparons ici les méthodes de tatouage sur des supports séparés et sur des supports emboîtés. Dans le cas où $n = 2$, les supports glissants sont équivalents aux supports emboîtés.

3.1 Robustesse au filtrage gaussien

Nous représentons le BER [4] en fonction de l'écart-type du filtre gaussien (*cf.* Figure 6). Nous remarquons trivialement que l'extraction dans les hautes fréquences est moins robuste que dans les basses, ce qui correspond à nos attentes. De ce point de vue, notons que les résultats ne sont pas perturbés sur les basses fréquences, la méthode des supports emboîtés reste équivalente à un tatouage sur des supports séparés.

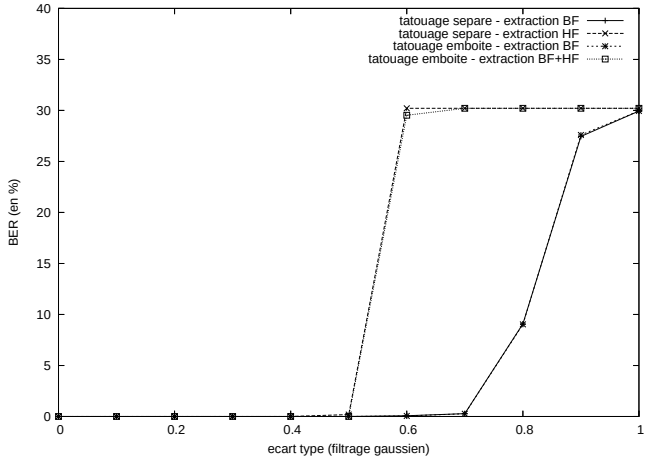


Figure 6 – *Représentation graphique du BER en fonction de l'écart-type lors du filtrage gaussien.*

3.2 Robustesse aux attaques valumétrique

De la même façon, nous représentons le taux d'erreur binaire en fonction du facteur valumétrique (*cf.* Figure 7). Le comportement de chaque support est identique. Autour de la valeur 1, il n'y a aucune perte des données. En faisant du *down-scaling* le taux d'erreur binaire croît brutalement pour se stabiliser à environ 30%. Pour l'opération inverse, la perte est plus progressive et moins importante.

4. BER : Bit Error Rate $(= \frac{\text{Nombre de bits erronés}}{\text{Nombre de bits total}})$

Ces résultats sont typiques des approches basées quantification. Ils peuvent être améliorés par l'approche de Abrardo *et al.* [6] qui propose une modification de QIM qui permet d'améliorer la robustesse aux attaques valumétriques. Quel que soit le type d'insertion, nous constatons qu'à l'extraction, le BER est quasiment identique.

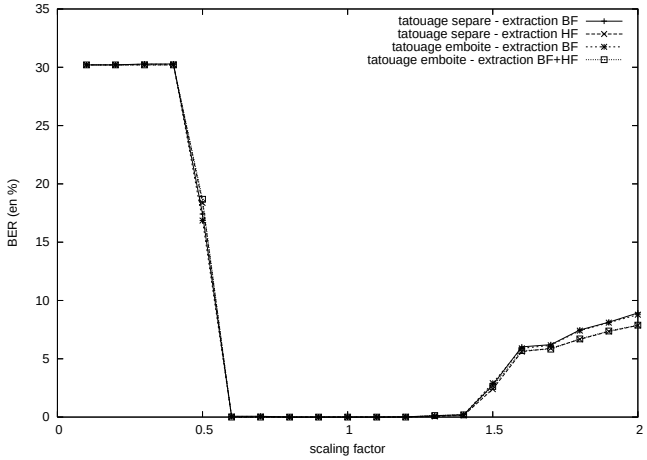


Figure 7 – *Représentation graphique du BER en fonction du facteur valumétrique.*

3.3 Robustesse à la compression JPEG

Nous nous intéressons ici à la compression JPEG. En fonction du facteur de qualité JPEG nous représentons le taux d'erreur binaire (*cf.* Figure 8). Egalement comme nous pouvions nous y attendre l'insertion dans les basses fréquences est plus robuste que dans les hautes.

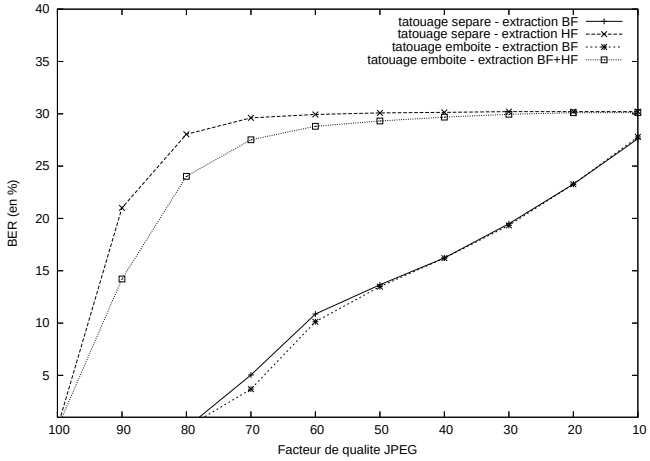


Figure 8 – *Représentation graphique du BER en fonction du facteur de qualité JPEG.*

Notons que la méthode de tatouage sur des supports emboîtés est plus robuste sur les hautes fréquences pour des facteurs de qualité $\geq 50\%$. Pour les facteurs de qualité $\leq 50\%$, la courbe suit asymptotiquement celle du tatouage sur des niveaux séparés en hautes fréquences. De plus, sur les basses fréquences, le chevauchement des supports d'insertion n'influe pas sur le comportement du taux d'erreur binaire en fonction du facteur de qualité.

4 Conclusion

Dans cet article, nous avons proposé de nouveaux supports d'insertion en vue d'un tatouage hiérarchique de message hiérarchique. Le message est découpé en k paquets $(k < n)$ et les plus importants sont tatoués dans les supports les plus robustes. Nous pouvons alors appliquer ces techniques sur des vidéos dans le cadre d'applications au suivi de copie et à de l'enrichissement.
L'approche des supports emboîtés possède un léger avantage par rapport à une technique de tatouage sur des supports séparés. Nous conservons des propriétés de robustesse équivalentes, et meilleures pour la compression JPEG, sans pour autant perturber les basses fréquences. Les approches QIM étant très rapides, il est envisagable de transposer ces approches à du tatouage de vidéos.

Références

[1] B. Chen et G.W. Wornell. Quantization Index Modulation : A Class of Provably Good Methods for Digital Watermarking and Information Embedding. *Information Theory, IEEE Transactions on*, 47(4) :1423–1443, 2001.

[2] H.J. Wang et C.C.J. Icuo. An Integrated Progressive Image Coding and Watermark System. *Acoustics, Speech and Signal Processing, Proceedings of the IEEE International Conference on*, 1998.

[3] A. Piper, R. Safavi-Naini, et A. Mertins. Coefficient Selection Methods for Scalable Spread Spectrum Watermarking. *Lecture Notes in Computer Science : Digital Watermarking : Second International Workshop*, pages 235–246, 2003.

[4] M. Antonini, M. Barlaud, P. Mathieu, et I. Daubechies. Image Coding Using Wavelet Transform. *Image Processing, IEEE Transactions on*, 1(2) :205–220, 1992.

[5] D. Taubman et M. Marcellin. *JPEG2000 Image Compression Fundamentals, Standards and Practice*. Kluwer Academic Publishers, 2002.

[6] A. Abrardo, M. Barni, F. Perez-Gonzalez, et C. Mosquera. Trellis-Coded Rational Dither Modulation for Digital Watermarking. *Lecture Notes in Computer Science*, 3710 :351, 2005.