

# Simulations Monte Carlo pour le tatouage robuste de séquence vidéo

Afef Chammem, Octavian Dumitru, Mihai Mitrea, Françoise Prêteux

*Institut TELECOM / TELECOM SudParis  
Département ARTEMIS*

9, rue Charles Fourier, 91011- Evry, France  
{afef.chammem, octavian.dumitru, mihai.mitrea, francoise.preteux@it-sudparis.eu}

**Résumé :** Aujourd'hui, la société du tout numérique doit maîtriser le foisonnement de contenus auto-produits lié à la multiplication des sources de production et des moyens de diffusion. L'enjeu se pose essentiellement en termes de préservation des droits de ces nouveaux auteurs puisque ceux-ci ne s'inscrivent pas dans une des chaînes traditionnelles de création. Les technologies par tatouage qui permettent d'insérer dans chaque œuvre numérique une information additionnelle de façon transparente et robuste aux attaques, apportent des solutions aux enjeux d'identification, de protection et de traçabilité des auteurs et des œuvres. L'état de l'art actuel est donné par les méthodes hybrides combinant étalement de spectre et information de bord. Toutefois, leur principal inconvénient réside dans leur temps de calcul prohibitif pour des services de vidéo à la demande et qui les rend inappropriées à une exploitation dans un contexte de temps réel. Cet article s'attaque à ce verrou. La contribution proposée est d'intégrer au schéma de tatouage hybride un module de simulation des attaques par générateurs Monte Carlo. L'étude des performances obtenues a montré un gain d'un facteur 100 tout en préservant les propriétés initiales de transparence et de robustesse du tatouage hybride. L'article discute en détail des analyses expérimentales conduites sur le corpus visuel établi dans le cadre du projet HD3D-IIO du Pôle de Compétitivité Cap Digital.

**Mots clés :** simulation Monte Carlo, tatouage hybride, étalement de spectre, information de bord, vidéo.

## 1 Introduction

Face au piratage numérique, le tatouage robuste offre une solution répondant aux enjeux de propriétés intellectuelles, d'authentification et de traçabilité [1-3].

Selon ce paradigme, une information additionnelle (*une marque*) est insérée dans une séquence vidéo, de façon imperceptible (*transparente*) pour l'observateur humain. La notion de transparence est liée à la visibilité des artefacts introduits lors de la procédure du tatouage. Bien que la transparence soit une notion intrinsèquement subjective, dépendant des aptitudes et du comportement de l'observateur humain, plusieurs mesures objectives ont été proposées [1-3].

Une méthode de tatouage est dite *robuste* dès lors que la marque peut être retrouvée aussi bien après les opérations usuelles de la vie normale d'un produit multimédia (*e.g.* compression avec pertes) qu'après des attaques malveillantes (*e.g.* StirMark). Ce point est tout à fait crucial au regard d'enjeux d'identification de droits de propriété, la marque permettant d'identifier tout comportement illicite.

La quantité d'information insérée (taille de la marque) dépend de l'application ciblée. Elle va d'un seul bit à des centaines ou milliers de bits (*e.g.* insertion d'un logo visuel dans une image).

Du point de vue théorique, tout schéma de tatouage peut être modélisé par un canal de transmission bruité (figure 1) [1-3]. L'information de copyright (logo, numéro de série...) cryptée avec la clé secrète représente un échantillon de la source d'information. Les éléments qui rendent difficile l'étape de détection peuvent être modélisés par un bruit de canal : il s'agit ici du document original, des traitements effectués et des attaques malveillantes. A noter qu'en pratique la marque est insérée dans une transformée de la séquence d'origine. La figure 1 correspond au choix de la transformée par ondelettes (TO). Si le document d'origine (non tatoué) n'intervient pas lors de la détection, la méthode est dite *aveugle*.

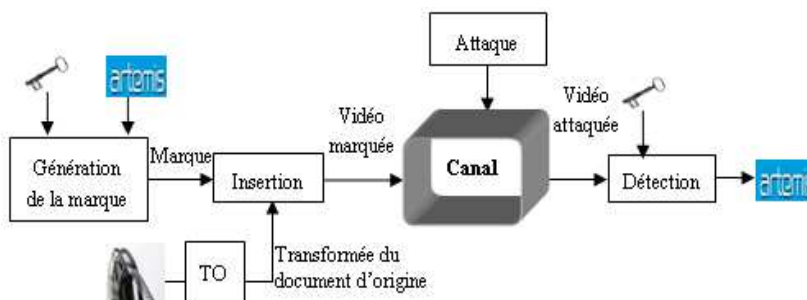


Figure 1 : Procédure de tatouage modélisée par un canal bruité.

Les deux principales techniques d'insertion s'appuient soit sur la théorie de la communication par spectre étalé (*spread-spectrum* - SS), soit sur la théorie dite par information de bord (*informed-embedding* - IE) [1].

Les techniques SS répartissent le spectre d'origine du signal-marqué dans des bandes de fréquences beaucoup plus larges que nécessaire selon le théorème d'échantillonnage de Shannon (*e.g.* 100 fois plus grandes). En pratique, ce type d'approche est très robuste aux attaques malveillantes, mais reste limité en termes de quantité d'information cachée. Les méthodes de tatouage IE, quant à elles, exploitent un autre théorème de Shannon qui exprime que tout bruit connu à la source, mais inconnu au détecteur, ne doit pas altérer les capacités du canal [4, 5]. L'approche IE intègre donc une certaine connaissance du contenu original dans la procédure d'insertion de la marque. En pratique, la quantité d'information insérée est très importante, mais la robustesse est faible.

L'équilibre opérationnel entre transparence, robustesse et quantité d'information insérée est atteint grâce aux approches hybrides, combinant principes du tatouage SS et IE [3,6]. Toutefois, leur principal inconvénient reste le temps calcul qui devient prohibitif pour une large classe d'applications, comme la VoD (vidéo à la demande) par exemple.

Au-delà des particularités d'implantation, comme toute méthode de communication, le tatouage reste limité par le bruit du canal. Il s'avère donc qu'aucune amélioration du tatouage n'est possible sans une connaissance mathématique des modèles du bruit. Une étude théorique menée au sein du Département ARTEMIS a d'une part mis en évidence que le modèle gaussien universellement utilisé pour représenter les effets de tout type d'attaque n'est pas toujours valide et d'autre part estimé le modèle de bruit réel [7, 8].

La contribution proposée par cet article est d'intégrer au schéma de tatouage hybride un module de simulation de type Monte Carlo pour les différents types d'attaque, pour surmonter le verrou lié à la vitesse d'exécution.

La méthode de tatouage hybride développée est tout d'abord décrite. Puis, le module de simulation des attaques par générateurs Monte Carlo est détaillé. Enfin, l'analyse comparée des nouvelles performances obtenues a été conduite sur le corpus audio-visuel établi dans le cadre du projet HD3D-IIO du Pôle de Compétitivité Cap Digital et les évaluations des performances en termes de transparence, robustesse et temps d'exécution discutées.

## 2 Méthode de tatouage hybride

Fondée sur les théories de la communication par spectre étalé et de l'information de bord, la méthode de tatouage hybride brevetée en 2007 par ARTEMIS et SFR [3, 6] a permis de trouver pour la première fois le compromis fonctionnel entre les contraintes antinomiques que sont la transparence, la robustesse et la quantité d'information insérée. Pour y parvenir, une nouvelle technique d'insertion qui s'appuie sur le comportement réel des attaques et non plus sur leur approximation gaussienne, a été développée (figure 2.a). La contre-partie en est un temps de calcul prohibitif pour certaines applications : par exemple, sur un ordinateur de type PC, doté d'un processeur Intel Centrino avec 1Go RAM, le tatouage d'une séquence de 100 trames prend plus d'une heure. L'analyse de la répartition moyenne des temps d'exécution entre les différents modules de la chaîne de tatouage (pré-traitement, insertion, post-traitement, détection), met en évidence que l'étape d'insertion compte pour 90% (figure 3), dont plus de 99% liés à prendre en compte le comportement réel des attaques. Une méthode efficace pour réduire ces derniers est de remplacer lors de l'insertion les attaques par leur simulation (figure 2.b). C'est l'objet de notre contribution dans cet article.

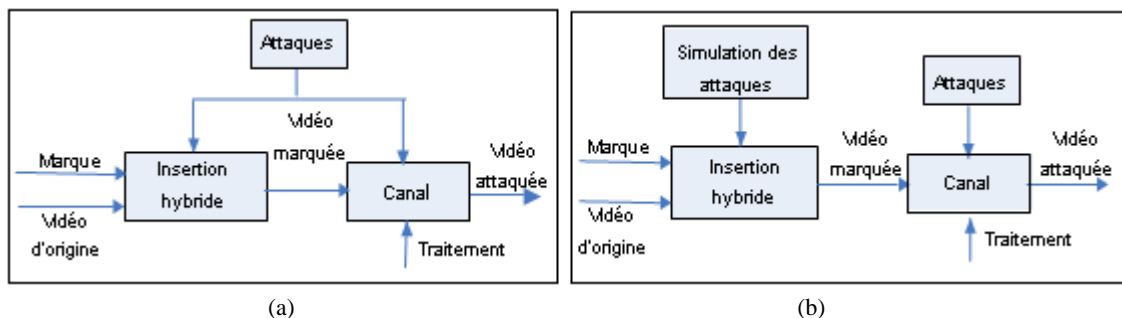


Figure 2 : Méthode de tatouage hybride : version initiale (a) et optimisée (b) par introduction d'un simulateur d'attaques.

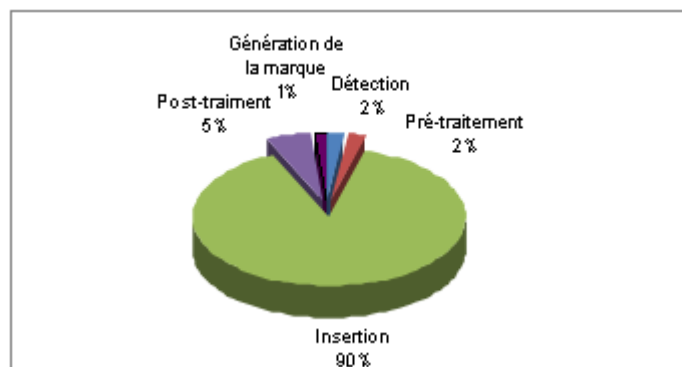


Figure 3 : Temps d'exécution pour les différentes opérations de la chaîne de tatouage.

### 3 Générateur Monte Carlo pour les attaques du tatouage

Une précédente étude [6] a démontré que le modèle de bruit additif gaussien n'est pas valide pour une large classe d'attaques malveillantes, *i.e.* la fonction de densité de probabilité gaussienne ne révèle pas le vrai comportement statistique des attaques. Elle a également fourni des modèles précis pour ces attaques.

Ces modèles sont exprimés sous forme de mélanges de gaussiennes permettant l'approximation non-paramétrique d'une densité de probabilité arbitraire [7][8].

Un mélange gaussien  $p(x)$  est la somme pondérée de densités gaussiennes :

$$p(x) = \sum_{k=1}^K P(k)N(x; m_k, \sigma_k),$$

où  $N(x; m_k, \sigma_k)$  représente la loi gaussienne de moyenne  $m_k$  et d'écart-type  $\sigma_k$ , et  $P(k)$  la probabilité *a priori* que la valeur mesurée soit produite par la  $k^{\text{ème}}$  composante du mélange.

Pour simuler une loi décrite par un mélange gaussien, on doit d'abord simuler le choix aléatoire, avec un poids  $P(k)$ , d'une variable gaussienne, puis simuler cette même variable, *i.e.* échantillonner la variable aléatoire gaussienne de paramètres  $m_k$  et  $\sigma_k$ .

Du point de vue algorithmique, cela se traduit par une simulation en deux étapes :

**1<sup>ère</sup> étape: Sélectionner l'index  $k$  de la variable aléatoire gaussienne**

- générer une valeur  $\alpha \in (0, 1)$  à l'aide d'un générateur uniforme de nombres aléatoires ;
- si  $\alpha \leq P_1$ , choisir  $k = 1$  (la première variable dans le mélange);
- sinon, choisir l'index  $k$  remplissant la condition suivante :

$$\sum_{j=1}^{k-1} P_j < \alpha \leq \sum_{j=1}^k P_j .$$

**2<sup>ème</sup> étape: Echantillonner la loi gaussienne correspondant à l'index  $k$**

- générer  $x_1$  et  $x_2$ , deux valeurs uniformément distribuées entre 0 et 1 ;
- calculer  $y = \sigma_k \sqrt{-2 \ln x_1} \cos(2\pi x_2) + m_k$  .

### 4 Résultats expérimentaux

Les séquences vidéos considérées pour l'étude expérimentale représentent aussi bien des vidéos de bonne qualité, susceptibles d'être distribuées sur Internet (50 séquences de 400 trames de  $608 \times 256$  pixels, codées à 512 kbit/s) que des vidéos à très bas débit typiquement diffusées sur réseaux mobiles (50 séquences de 400 trames de  $192 \times 80$  pixels, codées à 64 kbit/s). La figure 4 présente quelques trames extraites de deux séquences du corpus.

La marque est insérée dans les coefficients de la transformée en ondelettes (9,7), après leur ordonnancement décroissant. Les transformées ont été appliquées à un niveau de résolution  $N_r = 4$  pour les séquences de bonne qualité et  $N_r = 3$  pour celles à bas débit. A chaque fois, la quantité d'information insérée est de 400 bits.



Figure 4 : Exemples de trames extraites de séquences haute et basse qualité du corpus d'étude.

La première analyse réalisée a porté sur la transparence. L'évaluation subjective a été conduite à partir d'un panel de 10 observateurs de différents âges et professions. Celui-ci a établi que la méthode développée satisfait à la propriété de *fidélité* : aucune différence visuelle significative n'a pas pu être identifiée entre les vidéos initiales et les vidéos tatouées. Pour une évaluation objective de la transparence, les mesures suivantes ont été calculées (cf. tableau 1) : *Universal Image Quality Index* (UIQI), *Image Fidelity* (IF) et *Peak Signal to Noise Ratio* (PSNR). En notant par  $S$  et  $\hat{S}$  les deux séquences vidéos à comparer, par  $N_f$  leur nombre de trames et par  $W$  et  $H$  leur largeur et hauteur (en pixels), ces mesures sont définies comme suit :

$$UIQI = \frac{1}{N_f} \sum_{k=1}^{N_f} \frac{4\sigma_{S\hat{S}_k} \mu_{S_k} \mu_{\hat{S}_k}}{(\sigma_{S_k}^2 + \sigma_{\hat{S}_k}^2)(\mu_{S_k}^2 + \mu_{\hat{S}_k}^2)}, \text{ où}$$

$$\mu_{S_k} = \frac{1}{W \cdot H} \sum_{i=1}^W \sum_{j=1}^H S_{i,j,k}, \quad \mu_{\hat{S}_k} = \frac{1}{W \cdot H} \sum_{i=1}^W \sum_{j=1}^H \hat{S}_{i,j,k}, \quad \sigma_{S_k}^2 = \frac{1}{W \cdot H - 1} \sum_{i=1}^W \sum_{j=1}^H (S_{i,j,k} - \mu_{S_k})^2,$$

$$\sigma_{\hat{S}_k}^2 = \frac{1}{W \cdot H - 1} \sum_{i=1}^W \sum_{j=1}^H (\hat{S}_{i,j,k} - \mu_{\hat{S}_k})^2, \quad \sigma_{S_k \hat{S}_k} = \frac{1}{W \cdot H - 1} \sum_{i=1}^W \sum_{j=1}^H (S_{i,j,k} - \mu_{S_k})(\hat{S}_{i,j,k} - \mu_{\hat{S}_k}).$$

$$IF(S, \hat{S}) = \frac{1}{N_f} \sum_{k=1}^{N_f} \left( 1 - \frac{\sum_{i=1}^W \sum_{j=1}^H (S_{i,j,k} - \hat{S}_{i,j,k})^2}{\sum_{i=1}^W \sum_{j=1}^H (S_{i,j,k}^2 + \hat{S}_{i,j,k}^2)} \right)$$

$$PSNR = 10 \log_{10} \left( \frac{255^2}{MSE} \right) \text{ où } MSE = \frac{1}{N_f W H} \sum_{k=1}^{N_f} \sum_{i=1, j=1}^{i=W, j=H} |S_k(i, j) - \hat{S}_k(i, j)|^2.$$

Le tableau 1 présente les valeurs calculées pour le tatouage d'une même séquence originale (arbitrairement choisie du corpus) et pour 10 séquences tatouées avec des marques différentes. Il faut noter qu'aussi bien les mesures par corrélation (UIQI) que celles par erreur quadratique moyenne (IF et PSNR) offrent une très bonne transparence.

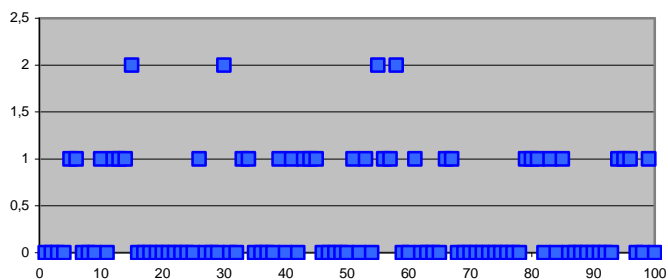
La seconde analyse conduite a porté sur l'évaluation de la robustesse. Nous voulions vérifier la résistance de la procédure de tatouage proposée aussi bien aux manipulations usuelles comme changement de format, compression avec perte, filtrage gaussien et transformations géométriques, ... qu'à des attaques malveillantes (e.g. StirMark). Dans tous les cas, la marque a été retrouvée avec

succès. Le tableau 1 montre le nombre d'erreur de détection (sur 400) pour les attaques des divers types, chaque attaque étant répétée 10 fois. La figure 5 ne considère que l'attaque StriMark, appliquée 100 fois : le nombre maximal d'erreur est de 2, pour 4% des cas.

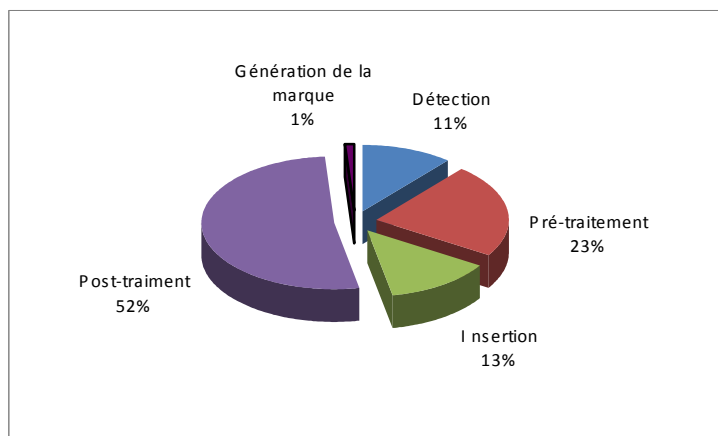
La dernière évaluation conduite (figure 6) a porté sur le temps de calcul de la nouvelle procédure.

**Tableau 1** : Evaluation quantitative des propriétés de transparence et robustesse.

		1	2	3	4	5	6	7	8	9	10	Moyenne
Transparence	UIQI	0,988	0,990	0,988	0,989	0,988	0,990	0,989	0,989	0,990	0,988	0,989
	IF	0,998	0,996	0,997	0,997	0,998	0,998	0,997	0,996	0,997	0,998	0,997
	PSNR	32,27	32,74	33,21	33,30	32,91	32,84	32,63	32,50	32,59	32,48	32,74
Robustesse	StirMark	0	1	0	0	1	1	0	0	1	0	0,4
	Filtrage gaussien	0	0	0	0	0	0	0	0	0	0	0
	Rehaussement	0	0	0	0	0	0	0	0	0	0	0
	Compression jpg	0	0	0	0	0	0	0	0	0	0	0
	Rotation +0.5	0	0	0	0	0	0	0	0	0	0	0
	Rotation -0.5	0	0	0	0	0	0	0	0	0	0	0



**Figure 5** : Nombre d'erreur pour 100 attaques StirMark.



**Figure 6** : Temps d'exécution des différentes opérations de la chaîne de tatouage.

Pour permettre une comparaison avec la figure 3, le même PC a été considéré (processeur Centrino avec 1 Go RAM). Les évaluations montrent tout d'abord un gain en temps d'exécution d'un facteur 100 par rapport à la méthode hybride de base. En effet, après avoir introduit le module de simulation d'attaques, la partie insertion ne représente plus que 13% du temps total requis par la chaîne de tatouage, en devenant ainsi comparable avec la partie détection (11%) et sept fois moins lourde que les parties de pré et post-traitement qui pèsent pour 75%.

## 5 Conclusion et perspectives

Cette contribution met en évidence l'intérêt d'introduire dans un schéma de tatouage hybride un module de simulation des attaques de type Monte Carlo pour prendre en compte le modèle réel de bruit. Il a été expérimentalement montré qu'un facteur 100 a été gagné lors de l'exécution de la méthode hybride de tatouage, tout en conservant les performances de transparence et robustesse de la méthode initiale. A souligner que d'un point de vue informatique, l'implantation est très simple et portable sous tout type d'environnement, fixe ou mobile, et tout type de terminal, y compris les clients légers.

Par rapport à la méthode de Miller [9], la méthode présentée dans ce papier garantit une meilleure transparence et robustesse, mais réduit la quantité d'information insérée.

Les perspectives de cette méthode porteront sur la possibilité de déployer des outils Monte Carlo pour l'optimisation de l'insertion informée et pour déduire par une approche statistique une technique optimale de détection.

## Références

- [1] I. Cox, M. Miller, J. Bloom, Digital watermarking, Morgan Kaufmann, 2002.
- [2] M. Mitrea, F. Preteux, "Tatouage robuste des contenus multimédias", in H. Chaouchi, M. Laurent-Maknavicius, La sécurité dans les réseaux sans fil et mobiles, Lavoisier, Paris, 2007.
- [3] M. Mitrea, S. Duta, F. Prêteux, "A unified approach to multimedia content watermarking", Proc. Third Taiwanese-French Conference on Information Technology (TFIT'2006), pp. 275-289, mars 2006.
- [4] C.E. Shannon. Channel with Side Information at the Transmitter. *IBM Journal*, 289-293, octobre 1958.
- [5] M.H.M. Costa. Writing on dirty paper. *IEEE transactions on Information theory*, IT-29(3): 439-441, mai 1983.
- [6] M. Mitrea, F. Prêteux, J. Nunez (for SFR and GET), "Procédé de tatouage d'une séquence video", French patent no. 05 54132 (29/12/2005), and EU patent no. 1804213 (04/07/2007).
- [7] M. Mitrea, O. Dumitru, F. Prêteux, A. Vlad, "Zero Memory Information Sources Approximating to Video Watermarking Attacks", Lecture Notes in Computer Science, Vol. 4705, pp. 409-423, août 2007.
- [8] O. Dumitru, M. Mitrea, Françoise Preteux, "Video Modelling in the DWT Domain", Proc. SPIE, Vol. 7000, pp. 7000, avril 2008.
- [9] L. Miller, G. Doërr, I. Cox. Applying Informed Coding and Embedding to Design a Robust High-Capacity Watermark. *IEEE transactions on image processing*, Vol. 13, No. 6, juin 2004.