

Tatouage vidéo par marquage hiérarchique des vecteurs mouvements.

BODO YANN¹ LAURENT NATHALIE¹ DUGELAY JEAN-LUC²

¹ France Telecom R&D 4, rue du Clos Courtel, BP91226
35510 CESSON-SEVIGNE, France

² Institut EURECOM, Multimedia Communication Dept
2229, Route des Cretes-B.P. 193
F-06904 Sophia Antipolis Cedex, France

Résumé

L'objet de cet article est de présenter une méthode de tatouage vidéo robuste contre des traitements classiques en vidéo (filtrage, compression, ...). Pour ce faire notre méthode repose sur l'élaboration d'une nouvelle approche de marquage hiérarchique des vecteurs de mouvement d'une vidéo.

Mots clefs

tatouage, vidéo, vecteurs mouvement.

1. Introduction

Avec l'essor récent des moyens de communications et notamment celui de l'Internet haut-débit, les échanges de données entre les personnes se sont généralisés et se sont intensifiés. Dans ce contexte il apparaît nécessaire de pouvoir contrôler les contenus multimédia payants. De nombreux systèmes sont apparus au cours de ces dernières années, il peut s'agir de systèmes de contrôle d'accès ou de systèmes de chiffrement, mais jusqu'à présent les différentes protections ne sont pas inviolable. Le plus souvent, la sécurité repose sur une clef qui doit rester secrète (clef qui finit souvent par être dévoilée comme dans le cas du système de protection CSS pour les DVD [6]). Le tatouage apparaît comme étant un élément complémentaire et bien adaptée à la protection de contenus multimédia. La protection peut se faire a posteriori contrairement aux systèmes classiques. Le principe est de cacher une information au sein d'un média de façon invisible et robuste. Ses applications sont nombreuses, il peut être utilisé dans des applications d'auteurs, de fingerprinting, de « smart content » etc. Pour plus de détails sur le tatouage, l'article [1] présente une bonne introduction, Tout comme l'article [2] qui lui propose un formalisme mathématique pour définir le problème du tatouage. Le système que nous avons développé s'inscrit dans un contexte de copyright, il s'agit d'un système de tatouage semi-aveugle, la vidéo originale n'est pas nécessaire mais il faut la marque originale. Le détecteur nous donne donc une réponse binaire.

2. Description technique

Un flux vidéo est composé d'une succession d'images. Le tatouage des séquences vidéos peut donc être vu comme une extension du tatouage appliqué à une série d'images fixes. En effet, il est possible de tatouer uniquement les images intra de la séquence. De nombreux schémas développés pour les images fixes peuvent alors être appliqués aux séquences vidéos.

Cependant les séquences vidéo possèdent d'autres propriétés qui peuvent être exploitées pour l'insertion de la signature :

- la taille brute d'une séquence vidéo est beaucoup plus importante que la taille d'une image fixe. L'espace d'insertion de la signature en est considérablement augmenté.
- la dimension temporelle du signal traité peut être utilisée pour l'insertion de la signature. Elle peut par exemple être insérée dans le mouvement des différents objets de la séquence.

Les séquences vidéo présentent aussi des contraintes différentes de celles des images fixes :

- la complexité du schéma de tatouage doit être faible. L'insertion et la détection de la signature doivent pouvoir s'effectuer à la volée dans la plupart des applications vidéo.
- le mouvement des objets augmente souvent la visibilité de la marque : ainsi une marque « fixe » ajoutée sur un objet qui se déplace sera d'avantage perceptible que si l'objet reste fixe.
- le flux vidéo est souvent compressé afin de réduire la taille brute des séquences. Le standard de compression le plus couramment utilisé est celui de MPEG2. L'insertion de la signature peut alors directement s'effectuer sur le format lui-même. Enfin l'insertion sur le format décompressé ne doit pas entraîner après compression, une augmentation de la taille du fichier.
- la présence de la signature dans une séquence vidéo peut permettre d'autres attaques que celles liées aux images fixes. Si la signature est redondante dans la séquence elle peut être estimée en calculant la moyenne des différentes images de la séquence. De plus la signature doit pouvoir être détectée après une perte de synchronisation produite par la sélection d'une

sous-séquence extraite ou la perte d'images de la séquence.

De ce fait, on peut classer l'ensemble des solutions de tatouage vidéo en 2 groupes :

- les techniques issues de l'image fixe,
- les techniques propres à la vidéo.

Notre approche s'inscrit dans la 2^{ème} catégorie.

2.1 Insertion

La marque est un mot binaire W constitué de 8 bits. Dans un premier temps nous ne réalisons aucun traitement sur cette marque, pour plus de robustesse, celle-ci pourra subir des traitements du type code correcteur d'erreur ou étalement de spectre. La phase suivante consiste à réaliser une estimation de mouvement sur la vidéo, les vecteurs ainsi obtenus permettent de créer une hiérarchie de vecteurs (phase de remontée) comme décrit sur la figure 1.

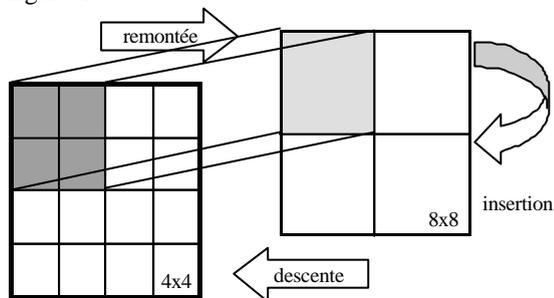


Figure 1

Parmi les vecteurs du niveau le plus élevé, on extrait un sous-ensemble de vecteurs, cette sélection peut se faire soit pseudo-aléatoirement soit de manière déterministe, soit à partir d'un masque psychovisuel.

La règle d'insertion est établie sur les vecteurs sélectionnés ci-dessus. Elle est basée sur une grille de référence (cf. Figure 2). Cette grille rectangulaire est générée dans le domaine cartésien (l'axe des abscisses correspond aux lignes des images et l'axe des ordonnées aux colonnes). La taille de ses éléments est $H \times K$. Avec $H, K=1, \dots, J$ (dans un mode particulier, on utilise une grille carrée avec $H = K = 7$)

Chaque vecteur mouvement calculé lors du processus d'estimation de mouvement (phase de remontée), est placé sur la grille. On détermine son point de référence qui correspond au point A de l'élément rectangulaire E ($Z1 \cup Z2$) auquel appartient le vecteur. Pour calculer le point de référence d'un vecteur de mouvement, on choisit le point de l'élément rectangulaire E qui minimise une distance entre C et un des 4 points de E (distance euclidienne par exemple).

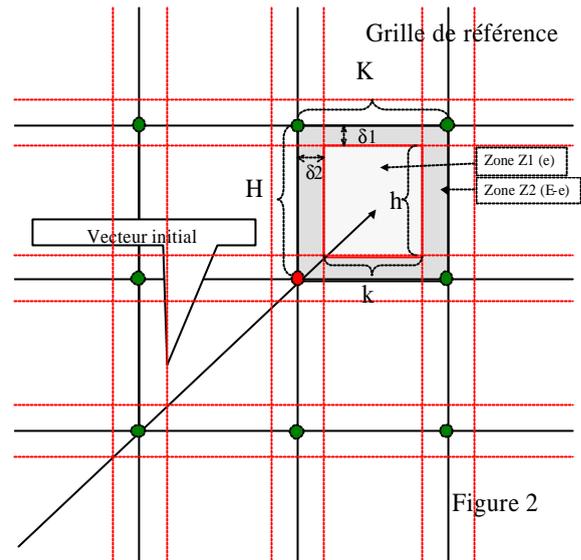


Figure 2

On modifie ensuite le vecteur en question suivant la règle suivante :

On génère un élément rectangulaire e inclus dans E de taille $h \times k$ avec $h = H - \delta 1$ et $k = K - \delta 2$ avec $\delta 1$ et $\delta 2$ tels que l'aire de e soit égale à l'aire de $E-e$. L'élément e correspond à la zone $Z1$ et l'élément $E-e$ à la zone $Z2$ (cf. Figure 2). On insère le bit de la façon suivante :

- Si on souhaite insérer le bit 1 et que le vecteur de mouvement est en $Z1$, on ne fait rien.
- Si on souhaite insérer le bit 0 et que le vecteur de mouvement est en $Z2$, on le déplace en $Z1$. Pour ce faire, on déplace C en D selon une direction axiale (cf. figure 3a):

$$\begin{cases} V'x = Vx + d_2 - PC_x - 1 \\ V'y = Vy + d_1 - PC_y - 1 \end{cases}$$

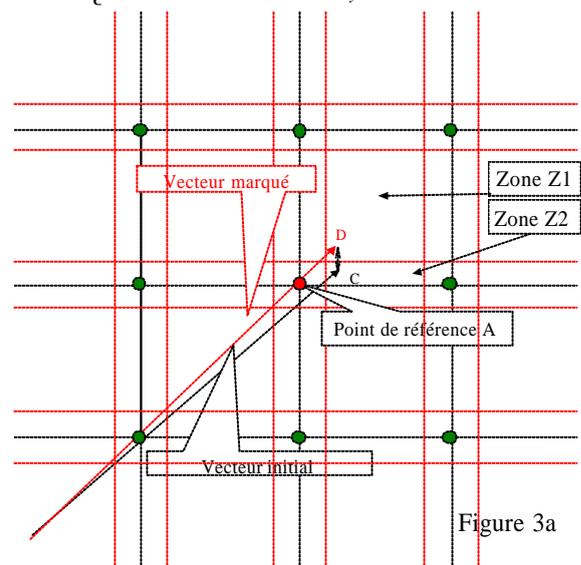


Figure 3a

ou selon une symétrie centrale de centre B (cf. figure 3b) :

$$\begin{cases} V'x = Vx + d_2 - PC_x - 1 \\ V'y = Vy + d_1 - PC_y - 1 \end{cases}$$

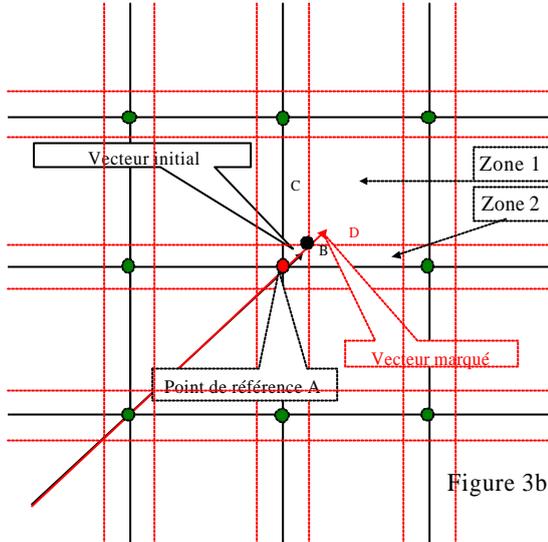


Figure 3b

- Si on souhaite insérer le bit -1 et que le vecteur de mouvement est en Z1, on le déplace en Z2 en réalisant une symétrie centrale pondérée de centre B. On déplace donc C en D. Dans un mode particulier de réalisation la pondération utilisée nous amène à la modification suivante (cf. figure 4) :

$$\begin{cases} V'x = Vx + d_2 - PC_x - 1 \\ V'y = Vy + d_1 - PC_y - 1 \end{cases}$$

- Si on souhaite insérer le bit -1 et que le vecteur de mouvement est en Z2, on ne fait rien.

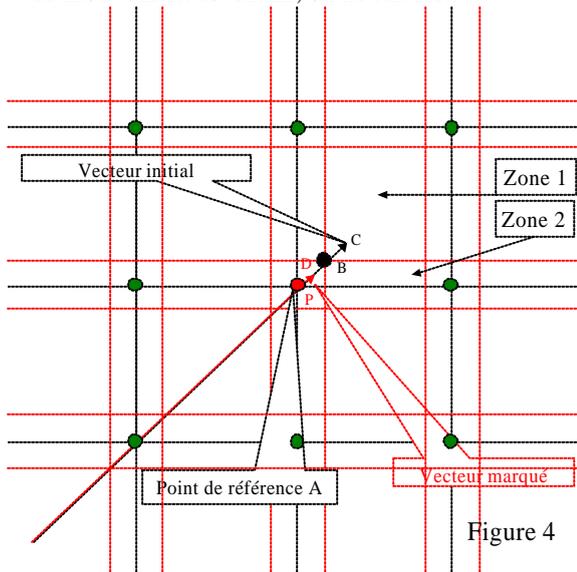


Figure 4

Les modifications ainsi réalisées sur le niveau le plus élevé sont ensuite transmises sur le niveau le plus bas (phase de redescente). Cette approche permet ainsi de créer une redondance dans l'insertion de la marque et de rendre le marquage plus robuste.

Enfin on effectue une compensation de mouvement qui peut être réalisée soit sur

l'ensemble des blocs soit sur les blocs sélectionnés et compléter l'image par les blocs originaux restants.

2.2 Extraction

La phase d'extraction est duale de la phase d'insertion. On extrait ainsi une marque qui est ensuite corrélée avec la marque originale afin d'établir un score de corrélation qui permet de déterminer la présence ou la non présence de la marque. Les règles couramment utilisées sont présentées en figure 5, nous avons utilisé la règle 2. La réponse du module d'extraction est donc une réponse binaire.

Règle de corrélation 1:
$$C = \frac{(\bar{W} - \bar{W}') * (W - \bar{W}')}{\sqrt{(\bar{W} - \bar{W}')^2 * (W - \bar{W}')^2}}$$

\bar{W}' : marque extraite

\bar{W} : moyenne de \bar{W}

W : marque originale

\bar{W} : moyenne de W

Règle de corrélation 2:
$$C_n = \frac{C_{n-1} * (n-1) + \left(1 - \frac{d(\bar{W}, W)}{8}\right)}{n}$$

n : nombre d'images

$d(\bar{W}, W)$: distance de Hamming

Figure 5

3. Résultats

Les tests ont été effectués sur 2 séquences différentes, les résultats sont visibles sur la figure 6a pour la séquence « tigre&dragon » comportant 134 images et la figure 6b pour la séquence « lord of the ring » comportant 463 images. Ces séquences sont au format YUV, leur taille est de 288*352 (format CIF).

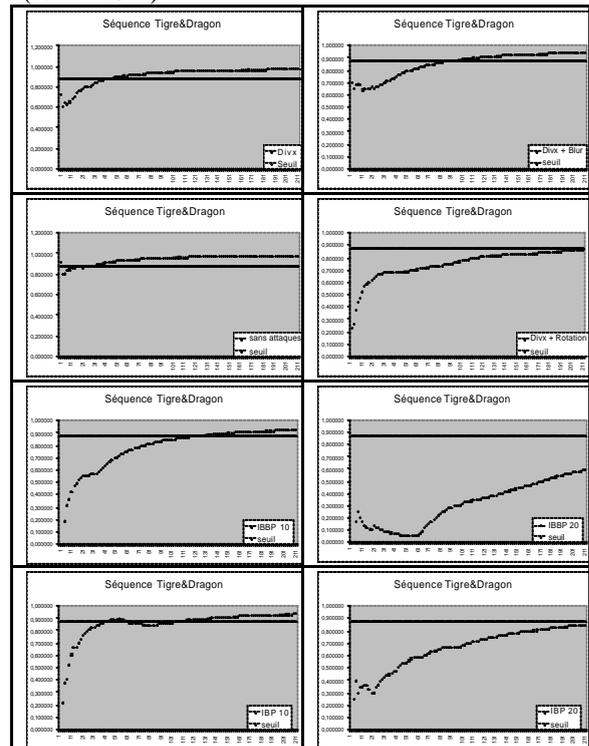


Figure 6a (score de corrélation accumulé, séquence Tigre&dragon)

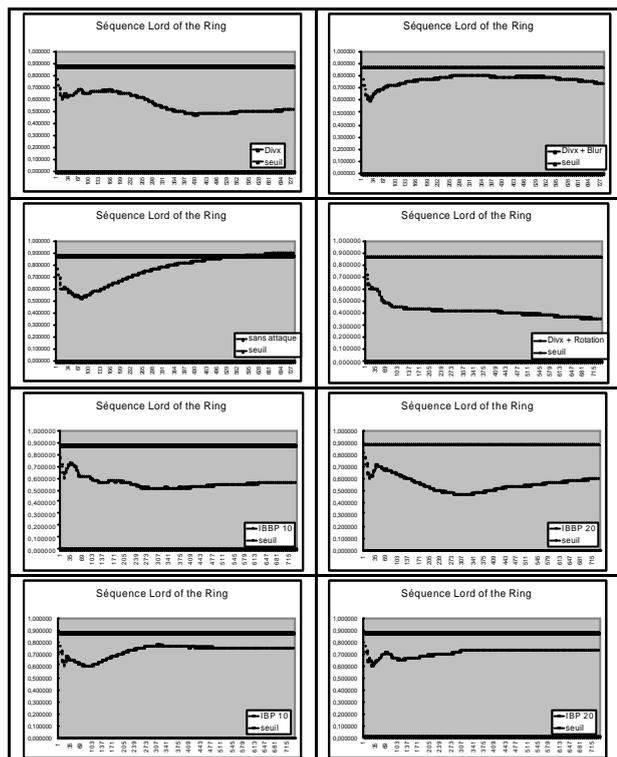


Figure 6b (score de corrélation accumulé, séquence Lord of the ring)

les résultats montre une bonne robustesse vis à vis différentes attaques que l'on peut considérer comme étant des traitements classiques en vidéo : compression (divx3 et h261), filtrage. Les taux de compression utilisés sont reportés sur la figure 7a et 7b.

Lord of the ring (taille originale : 137524 Ko)	Taille compressée (en Ko)
IBBP10	8841
IBBP20	2797
IBP10	8775
IBP20	2736
Divx3	1998

Figure 7a

Tigre & Dragon (taille originale : 39804 Ko)	Taille compressée (en Ko)
IBBP10	2223
IBBP20	618
IBP10	2214
IBP20	636
Divx3	522

Figure 7b

La robustesse dépend de la séquence marquée, on peut remarquer que la séquence « lord of the ring » présente quelques problèmes, c'est une séquence composée de plan très court ce qui crée une instabilité du marquage contrairement à la séquence « tigre & dragon » qui n'est constituée que d'un seul plan. Dans certains cas, au vu de l'allure de la courbe de corrélation, on peut supposer que le détecteur fini par converger vers la bonne marque mais le nombre d'image n'est pas suffisant. En outre, plus le détecteur mettra de temps à converger

vers la bonne marque, et plus la courbe de corrélation mettra de temps à dépasser le seuil fixé qui est de 87,5%.

Conclusion et perspectives

Notre approche s'avère robuste à des traitements que l'on rencontre couramment en vidéo. Il est toutefois nécessaire de poursuivre les tests sur d'autres séquences ainsi que sur d'autres manipulations et attaques. En outre il faut améliorer le critère de détection qui est trop sévère, en effet il est parfois possible que le détecteur converge vers la bonne marque tout en ayant un score de corrélation qui n'atteint pas le seuil fixé. Nous avons présenté ici une première approche de marquage qui est robuste à de nombreux traitements classiques, cette robustesse peut encore être améliorée notamment par une meilleure construction de la marque (utilisation de technique de cryptographie pour la résistance aux attaques de type collusion ou utilisation de technique de traitement du signal comme les codes correcteurs d'erreurs). En outre il persiste encore certains artefacts visibles qui apparaissent plus ou moins selon la séquence marquée. Leur suppression est actuellement à l'étude. L'utilisation d'un masque psychovisuel devrait permettre de diminuer voire d'éliminer complètement ces défauts. L'invisibilité de la marque n'est pas évidente à obtenir. En effet il n'y a pas de corrélation évidente entre les variations appliquées sur les vecteurs mouvement et l'impact visuel que cela peut provoquer. Nous étudions actuellement ce point.

Références

- [1] S.Roche, J.L.Dugelay, "Introduction au tatouage d'images" Annales des Télécommunications, vol. 54, pp. 427-437, 1999.
- [2] I.Pitas G.Voyatzis, "The use of watermarks in the protection of digital multimedia products" *Proceedings of the IEEE*, vol. 87, pp. 1197-1207, 1999.
- [3] F.Jordan, M. Kutter, T. Ebrahimi, "Proposal of a watermarking technique for hiding/retrieving data in compressed and decompressed video" ISO/IEC document JTC1/SC29/WG11/ MPEG97/M2281, juillet 1997.
- [4] J. Zhang, J. Li, L. Zhang, "Video watermarking technique in motion vector" Proc. XIV Brazilian symposium on computer graphics and image processing, 15-18.10.2001 pp 179-182.
- [5] I.J. Cox, J.K., F.T. Leighton, T. Shamoon, "Secure spread spectrum watermarking for multimedia" *Transactions on Image Processing*, 1997. 6(12): p. 1673-1687.
- [6] <http://www.dvd-copy.com/>