Etude du schéma Scalaire de Costa dans un domaine indépendant.

ILHEM BENKARA MOSTEFA [1], SOFIANE BRACI [1], CLAUDE DELPHA [1], REMY BOYER [1], MOHAMMED KHAMADJA [2]

[1] Laboratoire des Signaux et Systèmes (CNRS, Supelec, Université Paris-Sud 11)
3, Rue Joliot Curie - 91192 Gif sur Yvette, France.
[2] Laboratoire de Traitement du Signal, Département d'électronique et de l'ingénieur,
Université Mentouri, Route Ain El Bey, 25 000 Constantine, Algérie.

[1] {ilhem.benkara ; sofiane.braci ; claude.delpha ; remy.boyer}@lss.supelec.fr,
[2] m_khamadja@yahoo.fr

Résumé – Nous présentons ici une étude complète des performances du schéma principal de tatouage à information adjacente (Schéma Scalaire de Costa) dans un domaine indépendant. Dans cette étude, nous utilisons les propriétés de l'Analyse en Composantes Indépendantes à l'insertion et à l'extraction du message. Le but de ce travail est de mettre en œuvre un schéma de tatouage avec information adjacente dans un domaine indépendant ayant une grande capacité, forte robustesse à l'ajout de bruit et niveau de sécurité raisonnable. Nous comparons les performances obtenues à celles des schémas existants aux propriétés prouvées. Parmi nos résultats, nous montrons que le tatouage avec information adjacente dans un domaine indépendant permet globalement d'améliorer les performances par rapport aux schémas existants. Par exemple, en étudiant le taux d'erreur binaire (BER), nous montrons que la robustesse est améliorée de 20dB (1 décade) lorsque le rapport marque à bruit (WNR) vaut 0dB. Nous montrons également pour ce même WNR=0dB que la capacité pour notre schéma de tatouage est nettement plus grande que pour les schémas existants. Pour finir, l'étude des fonctions de densité de probabilité (PDF) des signaux originaux et des signaux marqués nous permet de montrer que le niveau de sécurité du schéma proposé est satisfaisant grâce à l'utilisation de l'Analyse en Composantes Indépendantes (ICA).

Abstract - This paper presents a complete study of the main informed watermarking scheme (Scalar Costa Scheme) in independent domain. In our study, we use the properties of independent components analysis (ICA) at the insertion and the extraction of the message. The aim of this work consist in develop an informed watermarking scheme in independent domain with high-capacity, huge robustness and correct security level. We compare the obtained performances to those of existing schemes that have proven properties. We show that the watermarking with side information in independent domain allows improving the existing performances. For example, considering the bit error rate, we show that the robustness is improved by 20dB (1 decade) when the watermark to noise ratio (WNR) is 0dB. We also show for the same value WNR = 0dB that the capacity for our proposed watermarking scheme is higher than for the existing schemes. The study of probability density function (PDF) of the original signals and the marked ones allow us to show that the security level of the proposed scheme is correct by the use of Independent Component Analysis.

1 Introduction

Le domaine du tatouage numérique a fait l'objet de nombreuses études ces dernières années. Parmi celles-ci plusieurs approches sont à dénombrer notamment autour de méthodes exploitant l'information adjacente (Quantization Index Modulation : QIM [1], Scalar Costa Scheme : SCS [2], ...), mais aussi des méthodes de transformation telles que les ondelettes, l'étalement [2], ou encore l'Analyse en Composantes Indépendantes (ICA) [3]. L'idée d'appliquer l'ICA dans le cadre du tatouage a été mise en œuvre avec deux approches différentes. La première d'entre elle a consisté à chercher à extraire, voire estimer, grâce à une analyse en composante indépendante, le message initialement inséré. Dan Yu et al [4] montrent que pour cette approche, le signal hôte, la marque et la clé d'insertion sont considérés comme des sources indépendantes qui sont mélangées et ainsi l'utilisation de l'ICA à l'extraction permet de séparer les sources pour estimer le message. Dans la seconde approche, l'ICA est utilisée à l'insertion du message. L'idée directrice de cette approche a été proposée dans les travaux de Gonzàles-Serrano et al. [5] et se résume selon le schéma de principe de la Figure 1 suivante. Dans ce travail, ce principe a été appliqué au tatouage par étalement de spectre.

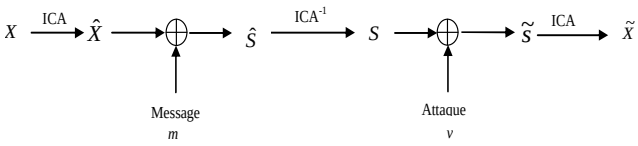

Figure 1 : Principe de tatouage dans un domaine indépendant

Dans ce schéma X est le signal hôte, $\hat{X}$ sont les sources indépendantes, m est le message à insérer, v le bruit d'attaque, $\hat{S}$ les composantes du signal marqué, S le signal marqué, $\tilde{S}$ le signal attaqué, et $\tilde{X}$ les sources marquées et attaquées. D'autres travaux se sont basés sur ce même principe pour améliorer les performances du tatouage par étalement de spectre, en associant à l'ICA une technique de tatouage basée sur

la quantification. Bounkong et al. [6] ont par exemple proposé l'association de l'ICA à la QIM à la fois à l'insertion et à l'extraction.
Notre travail est basé sur ce même schéma de principe (Figure 1), afin de proposer une méthode de tatouage informée s'appuyant sur le SCS, opérant dans un domaine transformé par ICA. Contrairement à Bounkong et al [6], dans notre approche, nous effectuons l'étude complète des performances de ce schéma de tatouage en termes de robustesse, capacité et sécurité selon la définition de Cachin [7]. De plus, nous comparons les performances obtenues aux différents systèmes référents connus pour leurs bonnes performances en termes de capacité : le SCS ; en terme de robustesse et de sécurité : le Spread Transform Scalar Costa Scheme (ST-SCS) [8]. Par le biais de cette étude, nous montrons qu'il est possible d'obtenir un très bon compromis de performances lorsque l'insertion et l'extraction du message se font dans un domaine indépendant. Nous validons les résultats obtenus grâce à une analyse théorique prouvant alors la possible utilisation de ce schéma pour des applications de tatouage robuste et invisible.

2 Schéma Scalaire de Costa dans un domaine Indépendant

Dans cette partie, nous présentons le schéma de tatouage utilisant l'ICA à l'insertion et à l'extraction du message noté SCS-ICA. L'ICA est une technique statistique dont l'objectif est de décomposer un vecteur $x \in \Re^m$ en une combinaison linéaire de sources indépendantes, c'est-à-dire $x = A \cdot \hat{x}$, où $\hat{x}$ est un signal à composantes indépendantes et A une matrice à coefficients réels notée matrice de mélange. Cette technique est appliquée à des problèmes où les sources peuvent être supposées indépendantes et pour lesquelles il est possible de trouver une matrice de séparation B telle que le vecteur obtenu par action de B sur x ait des composantes les plus indépendantes possibles [3].
La procédure de tatouage proposée dans un domaine indépendant basée sur le SCS comprend les principales étapes décrites sur la Figure 2.

1. L'algorithme ICA est appliqué au signal hôte x pour obtenir les composantes indépendantes $\hat{x}$.
2. Les sources obtenues $\hat{x}$ ont été quantifiées et tatouées, en utilisant le principe de la quantification scalaire du SCS [2] obtenir les composantes indépendantes tatouées $\hat{s}$ avec la marque w.
3. Une multiplication des composantes indépendantes tatouées par la matrice de mélange A (transformation ICA inverse), nous permet de récupérer les données tatouées s dans le domaine original du signal hôte.
4. Si on suppose que les données tatouées s sont transmises par l'intermédiaire d'un canal de communication pouvant introduire des distorsions qui seront modélisées par un bruit d'attaque v, le signal R reçu sera multiplié par la matrice de séparation B pour produire les composantes indépendantes tatouées et attaquées $\hat{R}$.
5. Le décodage du message se fait sans connaissance du signal hôte en utilisant le principe du SCS [2] pour extraire le message inséré $\hat{m}$.

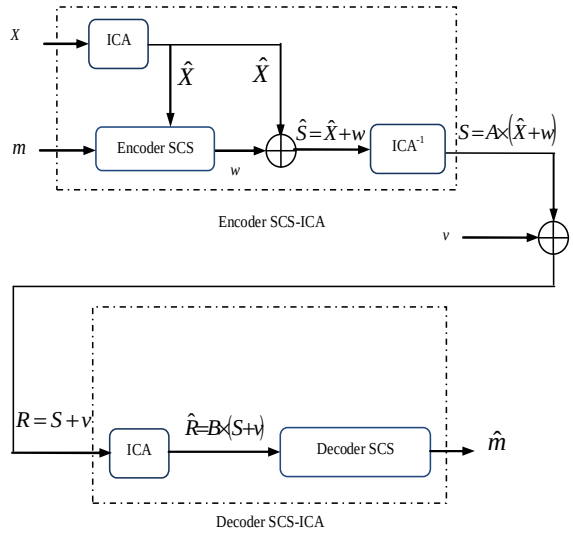

Figure 2 : Tatouage avec information adjacente dans un domaine indépendant.

La structure de ce schéma noté SCS-ICA étant largement inspirée du Spread Transform Scalar Costa Scheme (ST-SCS) proposé dans [8]. Pour ce schéma, la séquence pseudo aléatoire d'étalement est généralement utilisée en guise de clé secrète partagée entre l'encodeur et le décodeur. Dans le cas du SCS-ICA, nous procédons de façon similaire en considérant la matrice de séparation B comme clef secrète pour notre schéma basé sur l'ICA. Nous proposons également dans notre étude une évaluation comparative des performances entre ces schémas.

3 Résultats et discussions

Nous avons analysé les performances des schémas SCS, ST-SCS et SCS-ICA pour base d'image en niveau de gris de taille 512x512. Afin de s'assurer une imperceptibilité visuelle de la marque insérée, nous avons choisi d'utiliser un rapport de puissance entre le signal et la marque (DWR) suffisamment grand. Par exemple, nous utilisons un DWR=34,46 dB pour l'image Lena. Avant d'appliquer l'ICA, une transformation [5] a été effectuée sur l'image I de taille $N \times M$ comme suit pour obtenir le signal hôte X :

- L'image I est divisée en blocs $C_{p,q}$ de taille $k \times k$ dont les indexes ont pour domaine de variation $p = \{1,2,...,N/k\}$ et $q = \{1,2,...,M/k\}$. On obtient alors $C_{p,q}(i,j) = I(k(p-1)+i, k(q-1)+j)$, avec $i = \{1,2,...,k\}$ et $j = \{1,2,...,k\}$.
- Les blocs $C_{p,q}$ ont été transformés en vecteurs x_t^I où $t = (p-1)M/k + q$, tels que $t = \{1,...,MN/k^2\}$.

On a alors $x(k(i-1)+j)_{(M(p-1)/k+q)} = C_{p,q}(i,j)$. cette transformation est notée par : $\gamma(.)$ c'est-à-dire $x_t^I = \gamma(I,k)$.

Les lignes de x_t^I sont ensuite projetées en $l=k^2$ composantes indépendantes en utilisant l'algorithme FastICA [3] : $\hat{x}_t^I = B \cdot x_t^I$, où la matrice de séparation B, de taille $k^2 \times k^2$, est utilisée comme clef secrète. Ces composantes sont alors tatouées puis traitées selon le schéma de la Figure 2. Pour tous les résultats présentés, nous choisissons les conditions pour lesquelles le paramètre α de Costa est optimal [2].

Nous avons donc réalisé l'étude du taux d'erreur (BER) au décodage du message pour un niveau de bruit donné (WNR) de façon à évaluer la robustesse du schéma. Nous montrons, comme le prouve la Figure 3, que le schéma basé sur l'ICA a un niveau de robustesse très élevé. Par rapport au SCS il offre une diminution de plus d'une décade du BER pour un WNR=0dB. Par rapport au ST-SCS il offre également un BER bien meilleur tant que le facteur d'étalement τ reste inférieur à 5. Pour τ compris entre 5 et 8, le BER pour le SCS-ICA est meilleur que pour le ST-SCS dans la gamme ou le WNR est supérieur à 2dB.

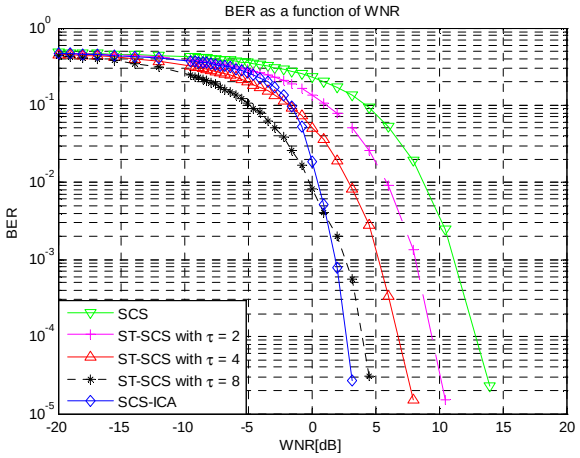

Figure 3 : Comparaison de la robustesse des différents schémas

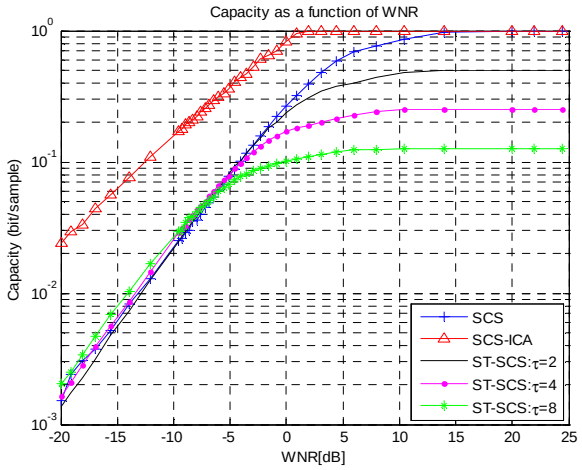

Figure 4 : Comparaison de la capacité des différents schémas

Nous avons ensuite analysé la quantité d'information que l'on peut décoder sans erreur pour un niveau de bruit donné (capacité). Dans le cas du ST-SCS il est bien connu que cette grandeur diminue lorsque le facteur d'étalement augmente [2]. Pour le SCS-ICA, la capacité obtenue est aussi bien plus grande, que pour les autres schéma, dans les zones ou le WNR est faible (WNR=[-20 à 5dB]) sans réduire le débit d'insertion. Pour les WNR plus élevés, on atteint la capacité théorique maximale fixée par le SCS contrairement au ST-SCS pour laquelle elle est divisée par le facteur d'étalement τ.

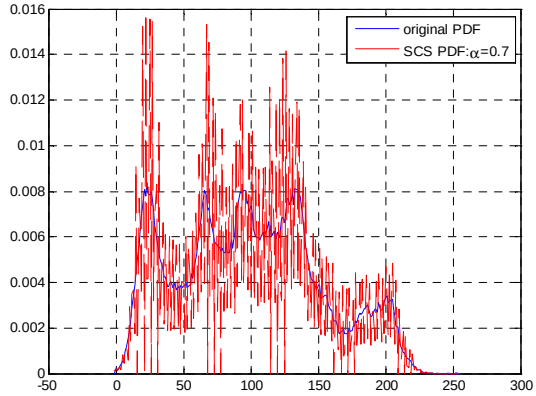

Figure 5 : Schéma SCS. PDF pour α=0.7

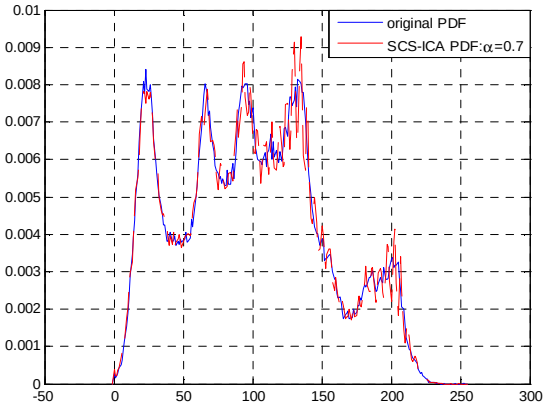

Figure 6 : Schéma SCS-ICA. PDF pour α=0.7

Nous avons également évalué la sécurité selon Cachin [7] pour le schéma proposé. Nous avons étudié la PDF des signaux marqués pour plusieurs valeurs du paramètre d'optimisation de robustesse (α), puis nous avons caractérisé les différences entre les PDF des signaux marqués et originaux grâce à la distance de Kullback-Leibler (KLD) [7]. Dans le cas du SCS, il est bien connu que de nombreuses discontinuités sont présentes sur la PDF du signal marqué dues à la quantification régulière (Figure 5). Il s'avère que ces celles-ci sont réduites lorsque $\alpha = 0.5$, ce qui impose d'utiliser le schéma dans des conditions sous optimales de robustesse et de capacité. Dans le cas du schéma avec l'ICA nous montrons que ces discontinuités sont réduites grâce à l'insertion sur les sources indépendantes (Figure 6). Par rapport au ST-SCS, nous montrons que ces performances en termes de sécurité sont moins bonnes mais restent satisfaisantes par rapport à celles du SCS (avec ou sans clef secrète) (Figure 7). Par rapport aux 3 schémas étudiés, le SCS-ICA est plus donc robuste, offre une meilleure capacité pour des débits

d'insertion identiques, améliore sensiblement la sécurité du SCS mais reste en deçà des performances de sécurité du ST-SCS.

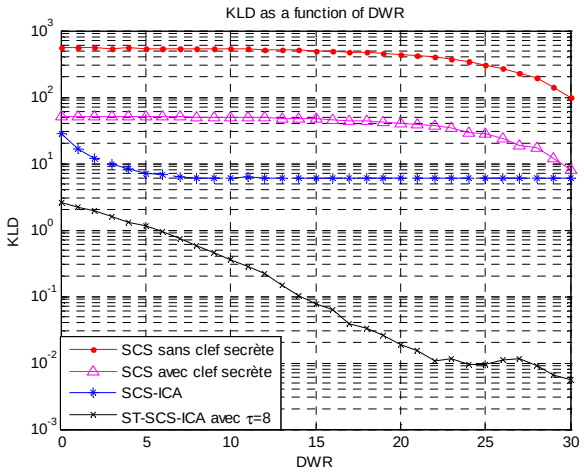

Figure 7 : Etude comparative de la KLD.

Pour mieux interpréter ces bonnes performances, nous avons étudié d'un point de vue théorique les modifications induites par l'ICA. Nous avons calculé les modifications sur le rapport marque à bruit, par rapport à celui du SCS, représentatif des transformations apportées sur la puissance de la marque.

Supposons que $S = X + A \cdot w$ est l'équation du processus d'insertion de la marque pour le SCS-ICA où: x est le signal hôte, A est la matrice de mélange et w la marque. Notons $w' = A \cdot w$ la marque pour le SCS-ICA et w la marque pour le SCS. Si on considère que la matrice A est composée de $a_{j'i'}$ éléments pour lesquels les indices j' et i' correspondent respectivement aux lignes et aux colonnes de la matrice avec pour domaines respectifs de variations $j' = \{1,..., k^2\}$ et $i' = \{1,..., k^2\}$. On peut alors écrire : $w'_{j'} = \sum_{i'=1}^{k^2} a_{j'i'} \cdot w_{i'}$ (1)

Où $w'_{j'}$ et $w_{i'}$ sont respectivement les composantes des vecteurs w' et w. Par conséquent, la variance $\sigma^2_{w'}$ de w' est donnée par :

$$\sigma^2_{w'j'} = E\left[w'^2_{j'}\right] = \sum_{i'=1}^{k^2} a^2_{j'i'} \cdot \sigma^2_{wi'} \quad (2)$$

où $E\left[w'^2_{j'}\right]$ est l'espérance de $w'^2_{j'}$. De plus $\sigma^2_{w'j'}$ et $\sigma^2_{wi'}$ sont respectivement les composantes des vecteurs de variances $\sigma^2_{w'}$ et σ^2_{w}.

Le calcul du rapport marque à bruit (WNR) pour chaque composante indépendante appliqué au SCS, considérant que σ^2_v est la variance du bruit, donne :

$$WNR_{(SCS)i'} = 10 \log_{10}\left(\frac{\sigma^2_{wi'}}{\sigma^2_v}\right) \quad (3)$$

L'équivalent pour le SCS-ICA est de la forme :

$$WNR_{(SCS-ICA)j'} = 10 \log_{10}\left(\frac{\sigma^2_{w'j'}}{\sigma^2_v}\right) \quad (4)$$

où $WNR_{(SCS)i'}$ et $WNR_{(SCS-ICA)j'}$ sont respectivement les composantes des vecteurs du WNR pour les schémas SCS et SCS-ICA. En combinant les équations 2, 3 et 4, on démontre que le WNR pour le SCS-ICA peut s'écrire sous la forme:

$$WNR_{(SCS-ICA)j'} = 10 \log_{10}\left(\sum_{i'=1}^{k^2} a^2_{j'i'} \cdot 10^{\frac{WNR_{(SCS)i'}}{10}}\right) \quad (5)$$

Compte tenu des propriétés de l'ICA, on aura les éléments $a^2_{j'i'} > 1$, donc $WNR_{(SCS-ICA)} > WNR_{(SCS)}$. Ainsi grâce à cette équation on note que les valeurs du WNR du schéma SCS sont amplifiées. Par conséquent, il y a une amplification que subit la puissance de la marque ce qui explique que pour un même niveau de bruit le taux d'erreur et la capacité sont meilleures.

4 Conclusion

Dans ce travail, nous avons analysé les performances d'un schéma de tatouage avec information adjacente basé sur l'insertion et l'extraction du message dans un domaine indépendant. Appliqué aux images, nous avons pu montrer que ce schéma (SCS-ICA) présente une meilleure robustesse et capacité que les schémas simples existants (SCS) et/ou qui emploient une transformée par étalement (ST-SCS) pour avoir de meilleures performances. Nous avons aussi montré que les performances que le niveau de sécurité d'un tel schéma est satisfaisant et meilleur que celui du SCS. Une analyse théorique a permis de confirmer les résultats obtenus.

5 Références

[1] B. Chen and G.W. Wornell. "Quantization index modulation", A class of provably good methods for digital watermarking and information embedding. *IEEE Trans. on Information Theory*, pages 1423–1443, 2001.

[2] J. J. Eggers, R. Bauml, R. Tzchoppe et B. Girod. "Scalar Costa scheme for information embedding", *IEEE Trans. on Signal Processing*, Avr. 2003.

[3] A. Hyvarinen. Survey on Independent Component Analysis. *Neural Computing Surveys*, Vol.2, pp.94-128, 1999.

[4] Dan Yu, Farook Sattar, and Kai-Kuang Ma, "Watermark detection and extraction using independent component analysis method," *EURASIP Journal on Applied Signal Processing*, vol. 1, pp. 92–104, 2002.

[5] F. J. González-Serrano, H. Y. Molina-Bulla, and J. J. Murillo- Fuentes, "Independent component analysis applied to digital image watermarking," in *Proc. of the IEEE ICASSP'01*, May 2001, vol. 3, pp. 1997–2000.

[6] S. Bounkong, B. Toch, D. Saad, and D. Lowe, "ICA for watermarking digital images," *Journal of Machine Learning Research*, vol. 4, no. 7-8, pp. 1471–1498, 2003.

[7] C. Cachin, "An Information-Theoretic Model for Steganography", in Proc. of Information Hiding, *Lecture Notes in Computer Science, Springer*, 1998.

[8] B. Chen and G. W. Wornell, "Achievable performance of digital watermarking systems," in *Proc. of the IEEE ICMCS '99*, June 1999, vol. 1, pp. 13–18.