TATOUAGE D'IMAGES BASÉ SUR LA TRANSFORMÉE DE FOURRIER DISCRÈTE

Gabriella Csurka, Frédéric Deguillaume, Joseph J. K. Ó Ruanaidh* et Thierry Pun
University of Geneva - CUI, 24 rue General Dufour,
CH 1211 Geneva 4, Switzerland
{Gabriela.Csurka,Frederic.Deguillaume,Thierry.Pun}@cui.unige.ch
http://cuiwww.unige.ch/~vision

1 INTRODUCTION

Le tatouage d'images est un champ de recherche récent et très prometteur, ayant pour but de décourager la copie et la distribution illicite de matériel protégé, ainsi que la protection de la propriété intellectuelle des données digitales. L'idée de base est d'insérer une information sous la forme d'un filigrane digital concernant le propriétaire de l'image, de même que des instructions concernant son copyright. Il est important que d'un côté l'extraction ou la suppression de cette information de l'image soit difficile voire impossible, et d'un autre côté que la distorsion introduite dans l'image par le filigrane soit minimale et invisible.

Beaucoup de techniques récentes de tatouage d'image sont inspirées des méthodes usuelles de codage et de compression. Dans ces méthodes, les informations sont ajoutées dans le domaine transformé de l'image, contrairement à d'autres techniques où le filigrane est inséré dans le domaine direct. Les transformées les plus utilisées sont la transformée cosinus discrète (TCD) [6, 1], la transformée de Fourrier discrète (TFD) [7], les ondelettes [6] et les fractales [10].

Afin de rendre la méthode de tatouage robuste, l'idée de base est de dissimuler la signature sous la forme d'un filigrane digital dans les composantes perceptuellement significatives de l'image, en tenant compte des propriétés du système visuel humain [6, 1, 9, 2]. Dans le domaine fréquentiel, les composantes perceptuellement significatives correspondent en général à des fréquences basses et moyennes. Cependant, si l'on modifie les basses fréquences de l'image, on observe un impact visuel important. Pour cette raison, il est important de choisir une bande de fréquences moyennes assurant un bon compromis entre la résistance maximale aux attaques et une dégradation de la qualité de l'image aussi peu visible que possible.

*Addresse actuelle: Siemens Corporate Research, 755 College Road East, Princeton, NJ 08540, US, oruanaidh@scr.siemens.com

Les techniques de “*spread spectrum*” sont souvent utilisées pour encoder les informations afin d'obtenir des filigranes digitaux pour le tatouage d'image [1, 7, 12]. L'idée de base de la technique *spread spectrum* est de transformer une séquence courte de bits (le message) en une séquence longue de telle manière que la séquence obtenue ressemble à du bruit. L'avantage principal du *spread spectrum*est qu'il est impossible de décoder le message encodé sans la connaissance de la clef privée utilisée pour l'encodage. Un autre avantage du *spread spectrum* est sa capacité à rejeter les interférences non désirées, dues à une transmission simultanée d'un autre message, ou à un attaque extérieure.

Pour ces raisons, l'approche présentée ici utilise un *spread spectrum* pour encoder le message contenant des informations concernant le propriétaire, ou un numéro de *hash-code* vers une table de hachage contenant ces informations. Pour générer la séquence *spread spectrum* l'on utilise des vecteurs binaires pseudo-aléatoires basés sur les m-séquences ou les codes de Gold. Le filigrane ainsi obtenu est ajouté dans le domaine de la transformée de Fourier discrète (TFD).

Afin de pouvoir détecter les transformations géométriques, un ensemble de positions du spectre de module de la transformée de Fourrier, que l'on appellera la grille de référence, sera modifiés de manière à y insérer des maxima locaux. Les positions de ces amplitudes dépendent de la clef du propriétaire. Ainsi, la grille de référence n'est connue au moment de l'extraction que si l'on possède cette clef. Pour décoder le message, les maxima locaux de la TFD sont extraits, et la transformation géométrique à calculer correspond à la transformation mettant en correspondance un maximum de points de la grille de référence avec les maxima locaux détectés. Grâce à ce calcul de transformation, la synchronisation du signal *spread spectrum* devient possible, et le filigrane peut être extrait et décodé. Cette technique *ne nécessite pas l'image originale*, ni pour estimer la transformation subie, ni pour extraire et décoder le filigrane.

Une autre contribution de cet article est l'utilisation d'une approche bayesienne pour calculer la probabilité qu'un filigrane ait été généré avec une clef donnée. L'avantage de cette approche est sa robustesse, car elle est capable de vérifier le propriétaire d'une image même si le message ne peut pas être correctement décodé, ceci en se basant sur la connaissance de la clef.

La sécurité du système est donc globalement basée sur la connaissance d'une clef appartenant au propriétaire, qui est nécessaire pour l'encodage, l'insertion, l'extraction et le décodage du filigrane. Un protocole assurant l'échange sécurisé des images et des clefs, et offrant une protection de copyright pour les images, est décrite dans [4]; cet aspect n'est pas traité ici.

2 LE FILIGRANE

Supposons que le message soit représenté sous forme binaire $\widehat{\boldsymbol{b}} = (\hat{b}_1, \hat{b}_2, ...\hat{b}_M)^\top$, où $\hat{b}_i \in \{0, 1\}$, et M est le nombre de bits du message à encoder. La séquence binaire est ensuite transformée pour obtenir le vecteur $\boldsymbol{b} = (b_1, b_2, ...b_M)^\top$ avec $b_i \in$

$\{1,-1\}$ en exploitant l'isomorphisme entre les groupes[1] ($\oplus$,{0,1}) et (*,{1,-1}). La correspondance $0 \rightarrow 1$ à $1 \rightarrow -1$ permet de remplacer le OU-exclusif par un produit durant le décodage. Définissant un vecteur binaire pseudo-aléatoire (en forme ± 1) $\boldsymbol{v}_i$ pour chaque bit b_i, le message peut être encodé ainsi: $\sum_{i=1}^{M} b_i \boldsymbol{v}_i = \mathbf{G}\boldsymbol{b} = \boldsymbol{w}$, où $\mathbf{G}$ est une matrice $N \times M$ avec la $i^{\text{ème}}$ colonne correspondant au vecteur $\boldsymbol{v}_i$ et le *spread spectrum* $\boldsymbol{w}$ est un vecteur de longueur N.

Pickholtz démontre dans [8] comment des vecteurs pseudo-aléatoires $\boldsymbol{v}_i$ peuvent être utilisés pour propager un signal donné afin de générer des séquences *spread spectrum*. Pour une bonne sécurité et une bonne résistance au bruit ces vecteurs pseudo-aléatoires doivent avoir des périodes longues afin de sembler parfaitement aléatoires. Ils doivent aussi être bien séparés en terme de corrélation. Les m-séquences et les codes de Gold vérifient bien ces propriétés [8, 3]. Les m-séquences sont des séquences de longueur maximale que l'on peut générer avec un registre LFSR (*linear feedback shift register*) d'une longueur donnée n [8]. Un tel registre génère des m-séquences avec une période $N = 2^n - 1$. La distribution spectrale de ces séquences ressemble à celle d'un bruit gaussien blanc. La corrélation entre une m-séquence et sa version déplacée vaut -1, et son autocorrélation est égale à N.

Grâce à ces propriétés, une possibilité de générer un *spread spectrum* $\boldsymbol{w}$ est d'utiliser une m-séquence $\boldsymbol{v}_1$ avec ses version décalées. Le vecteur $\boldsymbol{v}_{i+1}$ est alors obtenu à partir de $\boldsymbol{v}_i$ en décalant chaque élément de $\boldsymbol{v}_i$ d'une position à droite, le dernier élément de $\boldsymbol{v}_i$ devenant le premier élément de $\boldsymbol{v}_{i+1}$. Une autre possibilité de générer un *spread spectrum* $\boldsymbol{w}$ est d'utiliser une famille de codes de Gold [8, 3]. Une telle famille peut être obtenue à partir d'une m-séquence $\boldsymbol{v}_1$ grâce à une décimation q. Le vecteur $\boldsymbol{v}'_1$ de longueur N est obtenu à partir de $\boldsymbol{v}_1$ en considérant chaque $q^{\text{ème}}$ élément[2] de $\boldsymbol{v}_1$. Les vecteurs $\boldsymbol{v}_{i+1}$ de la famille des codes de Gold sont alors obtenus en multipliant élément par élément le vecteur $\boldsymbol{v}_1$ avec le $\boldsymbol{v}'_i$, où $\boldsymbol{v}'_{i+1}$ est obtenu à partir de $\boldsymbol{v}'_i$ en déplaçant chaque élément de $\boldsymbol{v}'_i$ d'une position à droite.

L'avantage des codes de Gold par rapport à une famille de m-séquences décalées est que pour un registre LFSR de longueur donnée on a d'avantage de choix[3] pour les $\boldsymbol{v}_i$. Un autre avantage des codes de Gold est qu'ils ont une meilleure propriété de corrélation dans le cas où seulement une partie de la séquence est considérée. Ceci peut avoir de l'importance si le filigrane est extrait partiellement.

Le filigrane $\boldsymbol{w}$ ainsi obtenu est inséré dans le spectre de module de la TFD. Pour assurer un bon compromis entre la résistance maximale aux attaques et une dégradation minimale de l'image, l'on fixe une bande dans les fréquences moyennes. La séquence *spread spectrum* est alors ajoutée aux valeurs des amplitudes de la bande choisie, la phase étant laissée inchangée. Pour une plus grande sécurité les positions des amplitudes affectées sont choisies en fonction de la clef. La TFD ainsi modifiée est finalement inversée pour obtenir l'image tatouée. La force du filigrane peut être choisie interactivement, ou bien calculée d'un manière adaptative en fonction de la moyenne et la variance des amplitudes dans la bande de fréquences choisie.

1. L' addition binaire modulo 2, $\oplus$ correspond à l'opération booléenne de OU-exclusif.
2. Pour obtenir un vecteur de longeur N, $\boldsymbol{v}_1$ est considéré plusieurs fois périodiquement.
3. Pour la famille de m-séquences, on dispose de la valeur initiale du registre LFSR déterminée par la clef, tandis que dans le cas des codes de Gold on peut également faire varier la valeur de q.

Pour extraire le filigrane, l'on considère le module de la TFD de l'image tatouée. Étant donné que les positions des coefficients de la bande de fréquences auxquelles les composants du filigrane ont été ajoutés sont connues, l'on peut extraire la séquence *spread spectrum* $\boldsymbol{w}' = \boldsymbol{w} + \boldsymbol{e}$, où $\boldsymbol{w}$ est le *spread spectrum* inséré et $\boldsymbol{e}$ un bruit additif. Le message $\boldsymbol{b}$ est alors décodé comme suit. Pour décoder le bit i du message $\boldsymbol{b}$ à partir de $\boldsymbol{w}'$, on calcule le produit scalaire (qui correspond à une corrélation) entre $\boldsymbol{w}' = \sum_{i=1}^{M} b_i \boldsymbol{v}_i + \boldsymbol{e}$ et $\boldsymbol{v}_i$:

$$B'_j = < \boldsymbol{w}', \boldsymbol{v}_j > = \sum_{i=1}^{M} b_i < \boldsymbol{v}_i, \boldsymbol{v}_j > + < \boldsymbol{e}, \boldsymbol{v}_j > \tag{1}$$

On a vu que pour les m-séquences et les Gold codes $< \boldsymbol{v}_i, \boldsymbol{v}_j > = -1$ pour $i \neq j$ et $< \boldsymbol{v}_i, \boldsymbol{v}_i > = N$. En les remplaçant dans (1), on trouve que: $B'_j = b_i N - (M-1) + < \boldsymbol{e}, \boldsymbol{v}_j >$. Mais, en général $M \ll N$, et la distribution de $\boldsymbol{e}$ peut être approchée par une distribution normale de moyenne égale à zéro. On a donc que $M-1$ et $< \boldsymbol{e}, \boldsymbol{v}_j >$ sont négligeables par rapport à N, ce qui donne $b'_j = sign(B'_j) = sign(b_j) = b_j$.

3 LA GRILLE DE RÉFÉRENCE

Si l'image a subi une modification géométrique telle qu'une rotation, un changement d'échelle, un découpage, etc, les positions des amplitudes marquées seront déplacées. Donc, avant d'extraire le filigrane, il est nécessaire d'estimer cette transformation géométrique afin de synchroniser le *spread spectrum* avant son décodage.

Afin de pouvoir détecter les transformations géométriques, l'on considère un ensemble de positions qui dépendent d'une clef du spectre de module de la TFD, appellé la grille de référence. Ces amplitudes sont modifiées de manière à devenir des maxima locaux. Pour extraire et décoder le message, l'on extrait d'abord les maxima locaux de la TFD. La transformation géométrique est estimée ensuite par mise en correspondance entre la grille de référence et l'ensemble des maxima locaux. Dans notre cas, l'on fait l'hypothèse que la transformation globale la plus générale qu'une image puisse subir est une transformation affine. Par conséquent, l'on utilise une paramétrisation affine (une transformation linéaire définie par une matrice 2×2 $\mathbf{T}$ plus une translation) pour représenter la transformation géométrique. En raison des propriétés de la TFD, cette transformation dans le domaine fréquentiel se décrit par quatre paramètres seulement (les coefficients de $\mathbf{T}^{-\top}$) car le module de la TFD est invariant par rapport à la translation. Il est donc suffisant d'avoir deux paires de correspondances entre les maxima locaux et la grille de référence pour estimer la transformation. La recherche s'effectue en considérant exhaustivement chaque paire de points de la grille avec chaque paire de maxima locaux, et en appliquant la transformation affine ainsi obtenue sur tous les points de la grille. La transformation retenue sera celle pour laquelle on a un nombre maximal de points de référence se trouvant au voisinage d'un maximum local après application de la transformation. Finalement, de manière à affiner la précision, la transformation affine finale est estimée sur toutes les correspondances trouvés en utilisant une méthode linéaire des moindres carrés. Grâce a cette transformation la synchronisation du signal *spread spectrum* est possible et le filigrane peut être correctement extrait et décodé.

4 L'APPROCHE BAYESIENNE

Considérons maintenant la séquence de bits extraite de l'image tatouée. Le but visé est évidemment de retrouver parfaitement la séquence insérée. Cependant, si la modification de l'image à été trop importante, il se peut que la séquence extraite diffère de quelques bits par rapport à celle insérée. La question qui se pose est alors de savoir quelle est la probabilité pour que le message extrait (même erroné) ait été généré avec une clef donnée.

Avant de répondre à cette question, remarquons que si nous avons une séquence binaire de m bits donnée, la probabilité qu'une séquence binaire aléatoire lui soit égale sur i bits est donnée par la distribution Bernoulli: $p(i) = C_m^i / 2^m$. Ceci nous montre qu'il est très improbable, par exemple, qu'une séquence aléatoire de 100 bits ait 80% de bits en commun avec une séquence donnée ($p(80) = 4.22 \times 10^{-10}$). Par contre, la probabilité que 50% des bits soient en commun est maximale ($p(80) = 0.0796$).

Cependant en général, lors de l'extraction l'on ne connaît pas le message inséré. On ne peut donc pas le mettre en correspondance avec les bits extraits. Par contre, la clef est connue et l'on peut supposer que la longueur du message l'est aussi[4]. Le message est codé à l'aide de m-séquences ou de codes de Gold pour donner le *spread spectrum* $\boldsymbol{w} = \mathbf{G}\boldsymbol{b}$, où $\mathbf{G}$ dépend seulement de la clef et de la longueur M du message. En remarquant que cette relation est linéaire et en assumant qu'à l'extraction l'erreur $\boldsymbol{e} = \boldsymbol{w}' - \mathbf{G}\boldsymbol{b}$ est gaussienne, l'on peut appliquer l'approche bayesienne décrite dans [11]. La probabilité pour que le *spread spectrum* $\boldsymbol{w}'$ extrait de l'image I' contienne un message de longueur M encodé avec la clef K est alors[5] :

$$p\left(K, M \mid \boldsymbol{w}', I'\right) \propto \frac{\pi^{-N/2}\,\Gamma\left(\frac{M}{2}\right)\Gamma\left(\frac{N-M}{2}\right)\left|\mathbf{G}^\top\mathbf{G}\right|^{-1/2}}{4\,R_\delta\,R_\sigma\left(\boldsymbol{b}'^\top\boldsymbol{b}'\right)^{M/2}\left(\boldsymbol{w}'^\top\boldsymbol{w}' - \boldsymbol{f}^\top\boldsymbol{f}\right)^{(N-M)/2}} \tag{2}$$

où $\boldsymbol{b}' = \left(\mathbf{G}^\top\mathbf{G}\right)^{-1}\mathbf{G}^\top\boldsymbol{w}'$ et $\boldsymbol{f} = \mathbf{G}^\top\boldsymbol{b}$ sont respectivement: le message estimé, et son ajustement au modèle $\mathbf{G}$. D'une façon similaire on peut calculer la probabilité qu'aucun message n'ait été inséré avec la clef K dans l'image I':

$$p(K,\, 0 \mid \boldsymbol{w}',\, I') \propto \frac{\pi^{-N/2}\Gamma\left(\frac{N}{2}\right)}{2R_\sigma(\boldsymbol{w}'^\top\boldsymbol{w}')^{N/2}} \tag{3}$$

Finalement, afin de décider si la clef k a été utilisée pour insérer un filigrane dans l'image I', nous calculons la probabilité logarithmique (log-probabilité) relative suivante:

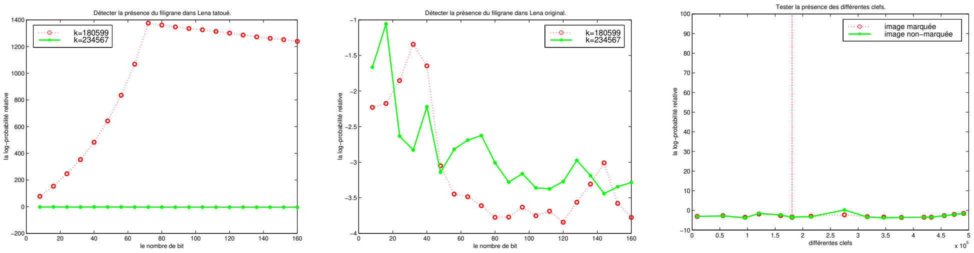
$$P_r = log\left(\frac{p(K,\, M \mid \boldsymbol{w}',\, I')}{p(K,\, 0 \mid \boldsymbol{w}',\, I')}\right) \tag{4}$$

Pour décider de la présence d'un filigrane appartenant au propriétaire de la clef K, nous comparons P_r avec un seuil 0, ou bien pour minimiser les fausses détections avec un seuil $S > 0$ (par exemple 1 ou 3).

4. Dans le cas où la longueur du message n'est pas connue, la log-probabilité relative est calculée pour les différents M possibles.

5. R_σ et R_δ sont des constantes et introduites comme facteurs de normalisation.

FIG. 1 – *Les images marquées utilisée pour les tests (Lena, Mandrill, Bateau).*



(a) (b) (c)

FIG. 2 – *Tests sur la longueur du message et sur differentes clefs.*

5 RÉSULTATS EXPÉRIMENTAUX

Pour l'expérimentation l'on a utilisé 3 images réelles. Le message de 72 bits est "CUI_Tests" et la clef 180599. Les images marquées sont présentées dans la figure 1.

Considérons tout d'abord le problème de la détermination de la longueur probable d'un filigrane utilisant cette clef par l'approche bayesienne. On va considérer une image tatouée (Lena) et estimer P_r pour différents longueurs. On peut remarquer que si l'on utilise la clef correcte on obtient une courbe qui atteint son maximum au nombre de bits du message inséré (72), par contre si l'on teste une mauvaise clef la courbe est pratiquement constante avec des valeurs négatives ou proches de zéro (Figure 2(a)). Si maintenant l'on répète la même expérience avec une image non marquée, on observe que les deux courbes ont le même comportement que la deuxième courbe du cas précédent (Figure 2(b)). En supposant connue la longueur du message, l'on teste la présence d'un filigrane avec différentes clefs dans l'image originale et dans celle tatouée. La figure 2(c) montre les log-probabilités relatives obtenues. Il est clair que la seule réponse positive ($P_r = 1375$) survient dans le cas de l'image tatouée et testée avec la clef correcte.

Supposons maintenant que l'on connaîsse la clef et la longueur du message. Pour tester le comportement de l'algorithme par rapport à différents attaques, l'on applique le programme Stirmark 3.1[6] de Fabien Petitcolas [5] aux images tatouées. Les résultats obtenus sont présentée dans le tableau 1. Dans la figure 3 nous présentons le comportement de la log-probabilité relative pour quelques uns de ces tests.

6. http://www.cl.cam.ac.uk/ fapp2/watermarking/stirmark/.

Tests	$ber = 0$	$ber \leq 5\%$	$ber > 5\%$	$P_r > 1$	$P_r \leq 1$
JPEG	24/36	4/36	8/36	33/36	3/36
Filtres ± JPEG	33/33	0/33	0/33	33/33	0/33
CE ± JPEG	32/36	0/36	4/36	35/36	1/36
CP ± JPEG	48/48	0/48	0/48	48/48	0/48
TC ± JPEG	85/96	5/96	6/96	94/96	2/96
TCA ± JPEG	87/96	3/96	6/96	94/96	2/96
Découpage ± JPEG	42/54	4/54	8/54	47/54	7/54
Affine ± JPEG	29/36	7/36	0/36	36/36	0/36
Shearing ± JPEG	17/36	19/36	0/36	36/36	0/36
LCS ± JPEG	30/30	0/30	0/30	30/30	0/30
Flip	3/3	0/3	0/3	3/3	0/3
Stirmark	0/3	0/3	3/3	0/3	3/3

TAB. 1 – *Les résultats des tests Stirmark 3.1. On a noté par ber (*bit error ratio*) le pourcentage de bit erroné et par ± JPEG le fait que la ligne contiennent les résultats pour les images modifiée non-compressée et compressée avec JPEG90. Les opérations de filtrage testées sont des filtres gaussiens, des filtres médians, le filtre laplacien et le Sharpening. Les variations des paramètres sont: les facteurs de qualité de JPEG entre 10 et 90; les changements d'échelle (CE) entre 0.5 et 2; les changements de proportion (CP) entre 0.8 à 1.2; les angles de rotations pour TC (image et tournée et coupée) et TCA (image tournée, coupée et agrandie) entre -2 et 90; la proportion coupée de l'image entre 1% et 75%; le pourcentage du cisaillement (shearing) horizontal et/ou vertical entre 1% à 5%; le nombre de lignes et/ou de colonnes supprimées (LCS) entre 1 et 17 avec un espacement régulier.*

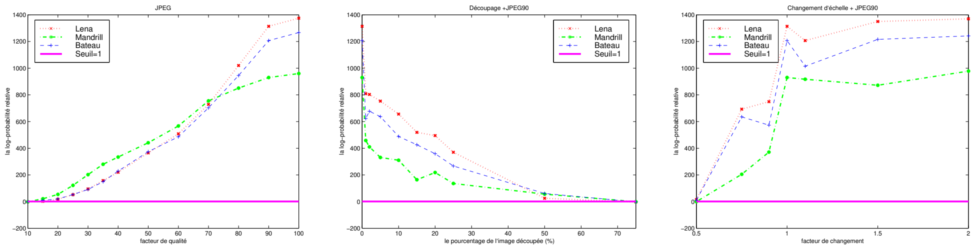
6 CONCLUSION

On a présentée ici une approche de tatouage d'images basée sur la transformée de Fourier discrète (TFD). Le filigrane ajouté au domaine de la TFD est obtenu en encodant le message à l'aide de séquences *spread spectrum* basées sur des vecteurs binaires pseudo-aléatoires telles que les m-séquences ou les codes de Gold. Afin de pouvoir détecter et compenser les transformations géométriques, l'on utilise une grille de référence selon laquelle les amplitudes de la transformée de Fourrier sont modifiées. Grâce à cette technique il est possible de synchroniser le signal *spread spectrum* et le filigrane peut être extrait et décodé. Afin d'être capable de vérifier le propriétaire d'une image même si le message n'a pas été correctement décodé, et en se basant sur la seule connaissance de la clef, l'on utilise une approche bayesienne pour calculer la probabilité qu'un filigrane ait été généré avec une clef donnée.

Les expériences effectuées montrent que cette méthode est résistante à la plupart des attaques, avec un impact visuel acceptable du filigrane sur la qualité de l'image.

Références

[1] I. Cox, J. Killian, T. Leighton, and T. Shamoon. Secure spread spectrum watermar-



(a) (b) (c)

FIG. 3 – *Les log-probabilités relatives pour quelques tests de Stirmark 3.1. (a) JPEG, (b) Découpage + JPEG90 (c) Changement d'échelle + JPEG90.*

king for images, audio and video. In *Proc. of the IEEE Int. Conf. on Image Processing ICIP-96*, pages 243–246, Lausanne, Switzerland, 1996.

[2] J. F. Delaigle, C. De Vleeschouwer, and B. Macq. Digital Watermarking. In *Conference 2659 - Optical Security and Counterfeit Deterrence Techniques*, San Jose, February 1996. SPIE Electronic Imaging : Science and Technology. pp. 99-110.

[3] E. H. Dinan and B. Jabbari. Spreading codes for direct sequence CDMA and wideband CDMA cellular network. *IEEE Communications Magazine*, pages 48–54, June 1998.

[4] Alexander Herrigel, Joe J. K. Ó Ruanaidh, H. Petersen, Shelby Pereira, and Thierry Pun. Secure copyright protection techniques for digital images. In *International Workshop on Information Hiding*, Portland, OR, USA, April 1998.

[5] M. Kutter and F. A. P. Petitcolas. A fair benchmark for image watermarking systems. In *Electronic Imaging '99, Security and Watermarking of Multimedia Contents*, volume 3657, pages 219–239, San Jose, CA, USA, January 1999.

[6] Joe J. K. Ó Ruanaidh, W. J. Dowling, and F. M. Boland. Watermarking digital images for copyright protection. *IEE Proceedings on Vision, Signal and Image Processing*, 143(4):250–256, August 1996.

[7] Joe J. K. Ó Ruanaidh and Thierry Pun. Rotation, scale and translation invariant spread spectrum digital image watermarking. *Signal Processing*, 66(3):303–317, 1998.

[8] R. L. Pickholtz, D. L. Schilling, and L. B. Milstein. Theory of spread spectrum communications – A tutorial. *IEEE Trans. on Communications*, COM-30(5):855–884, May 1982.

[9] I Pitas. A method for signature casting on digital images. In *Proc. of the IEEE Int. Conf. on Image Processing ICIP-96*, pages 215–218, Lausanne, Switzerland, 1996.

[10] J. Puate and F. Jordan. Using fractal compression scheme to embed a digital signature into an image. In *Proc. of SPIE Photonics East'96 Symposium*, November 1996.

[11] J. J. K. Ó Ruanaidh and W. J. Fitzgerald. *Numerical Bayesian Methods Applied to Signal Processing*. Series on Statistics and Computing. Springer-Verlag, 1996.

[12] A. Z. Tirkel, C.F. Osborne, and T.E. Hall. Image and watermark registration. *Signal processing*, 66:373–383, 1998.