Tatouage d'images avec des données biométriques révocables pour la preuve de propriété

Morgan Barbier (morgan.barbier@ensicaen.fr)*
Christophe Rosenberger (christophe.rosenberger@ensicaen.fr)*

Résumé : Dans ce papier, nous traitons la problématique de preuve de propriété d'une image afin de lutter contre la copie illicite de celle-ci. Deux techniques sont combinées à savoir le tatouage d'images et la biométrie révocable pour qu'une autorité de confiance puisse insérer une marque de propriété et la vérifier. Nous illustrons la robustesse de la méthode face à des altérations simulant des attaques volontaires ou involontaires de l'image tatouée.

Mots Clés : biométrie, tatouage d'images, biométrie révocable.

1 Introduction

Dans cette dernière décennie, beaucoup de travaux de recherche ont proposé différentes méthodes pour gérer les droits d'auteurs des images en utilisant des briques et des protocoles cryptographiques [Faz06, Moh09]. Cependant, ces schémas réalisent une vérification peu sécurisée du propriétaire de la donnée. En effet, ces systèmes peuvent difficilement lier l'identité de l'individu à ces droits d'utilisation. Afin de pallier ce problème, des chercheurs ont pensé à utiliser de la biométrie. L'insertion de données biométriques dans une image a été proposée pour la première fois en 2004 pour lier l'iris d'un photographe avec les photos qu'il a prises [BF04]. Cette même technique a également été proposée pour des applications de multibiométrie, consistant par exemple insérer les données biométriques du doigt dans une image de visage [VSN+06, KD07], mais aucunement à des fins de preuve de propriété de l'image. Le problème majeur de ces solutions est que ces méthodes insèrent des données sensibles dans une image distribuée, portant grandement atteinte à la vie privée de l'utilisateur. De nouveaux schémas ont été proposés afin de rendre une donnée biométrique révocable [TNG04]. La connaissance de cette donnée révocable ne permet pas de remonter à la donnée biométrique protégeant ainsi la vie privée des utilisateurs.

Dans cet article, nous proposons un nouveau schéma de preuve de propriété d'une image exploitant la biométrie révocable. Cela nous permet, à l'aide également de protocoles cryptographiques, d'assurer la sécurité du tatouage tout en respectant la vie privée de son propriétaire. Dans un premier temps, nous définissons les exigences que doit respecter ce type de schéma. Nous présentons par la suite le schéma proposé. Nous effectuons ensuite une analyse de la sécurité et de la protection de la vie vie privée et nous montrons les résultats de nos expérimentations illustrant l'apport de l'approche proposée.

*. ENSICAEN, Laboratoire GREYC, 17 rue Claude Bloch, 14000 Caen, France

2 Exigences de sécurité et de protection de la vie privée

Un système biométrique manipule, par définition, des données personnelles sensibles des utilisateurs. Ces données doivent être protégées pour éviter l'usurpation, leur modification ou falsification. Dans notre contexte, on définit trois acteurs principaux. On considère tout d'abord le **client** qui souhaite acquérir les droits d'utilisation d'un média proposé par un **fournisseur**. Lorsque le client souhaite revendiquer son droit d'utilisation, il contacte une **autorité** afin de prouver qu'il possède bien les droits associés au média (notamment pour prouver qu'il est bien le propriétaire de celui-ci). Nous introduisons les principales exigences de sécurité et de protection de la vie privée associées à ce genre de système :

- E_1 : La **preuve de propriété** assure à un client légitime qu'il peut prouver à chaque instant ses droits afin de démontrer que le média n'est pas une copie illicite.
- E_2 : La **non-associabilité** des marques dans différents médias possédés par un même client garantie qu'il n'est pas possible qu'une personne non autorisée puisse lier ces différentes marques à ce client.
- E_3 : La **confidentialité des données du client** assure que les données d'authentification du client ne soient connues ni par le fournisseur, ni par l'autorité.
- E_4 : La **souveraineté des données** implique que la vérification de la marque de propriété dans le média ne puisse être réalisée qu'avec le consentement et le contrôle du client.
- E_5 : La **non-falsification** empêche un client/attaquant de fabriquer une marque valide.
- E_6 : La **non-répudiation** empêche le fournisseur de nier la vente d'un média.

3 Méthode proposée

La méthode proposée consiste à insérer une marque dans une image à des fins de preuve de propriété par une autorité (voir Figure 1). La marque est calculée à partir d'une donnée biométrique révocable afin de permettre une vérification d'identité sécurisée et respectueuse de la vie privée. L'algorithme de BioHashing est utilisé pour calculer cette donnée révocable.

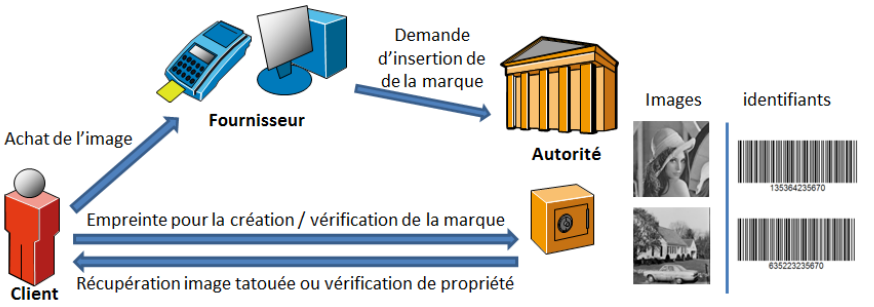

Figure 1: Schéma général de la méthode proposée

3.1 Algorithme de BioHashing

Nous supposons avoir une empreinte digitale de l'utilisateur décrite par un vecteur de paramètres appelé FingerCode. L'algorithme de BioHashing transforme ce vecteur à valeurs réelles de taille n en un vecteur binaire appelé BioCode de taille $m \leq n$ (voir Figure 2). Teoh *et al.* sont les premiers à l'avoir défini dans [TNG04].

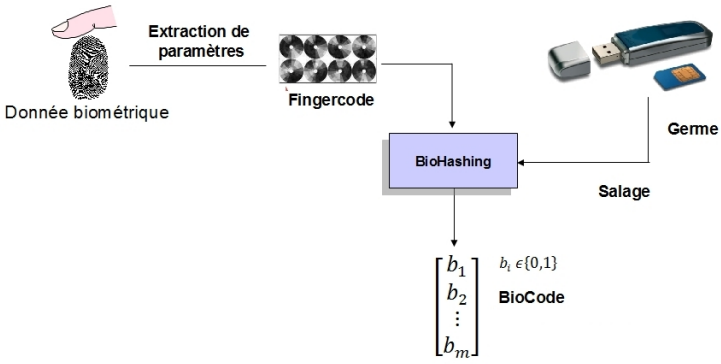

Figure 2: Schéma de BioHashing

Lorsqu'il est appliqué dans un contexte d'authentification, le *BioCode référent* (calculé à partir du FingerCode après l'enregistrement et après avoir présenté le germe considéré comme un secret) est comparé au *BioCode capturé* (calculé à partir du FingerCode obtenu après une nouvelle capture et le même secret) avec la distance de Hamming. Si cette valeur est inférieure au seuil déterminé par l'administrateur du système, l'identité de l'utilisateur est vérifiée. L'algorithme de BioHashing permet une transformation isométrique non inversible de la donnée biométrique. Le secret garantie des propriétés de diversité (différents BioCodes pour le même utilisateur) et de révocabilité (régénération du BioCode en cas d'interception).

3.2 Méthode de tatouage d'images

Nous utilisons dans cette étude une méthode de tatouage proposée par Wenyin et Shih [WS11]. Cette méthode utilise des paramètres de texture d'images appelés LBP (Local Binary Pattern) pour sélectionner les pixels à marquer. Cette approche a montré une bonne robustesse à différentes altérations classiques (compression, rognage...) simulant des attaques volontaires ou involontaires de l'image tatouée.

3.3 Insertion de la marque biométrique

Nous utilisons la méthode de tatouage précédente en utilisant une marque biométrique. La marque biométrique est définie par la répétition du BioCode du propriétaire de l'image. Dans notre cas, nous répétons 16 fois le BioCode de 256 bits (obtenu à partir d'un FingerCode de taille 512) pour obtenir une marque 64 x 64 pixels. Nous avons utilisé une

germe du générateur aléatoire calculée comme le hachage (avec la méthode SHA-256) d'un identifiant de l'image généré et connu seulement par l'autorité.

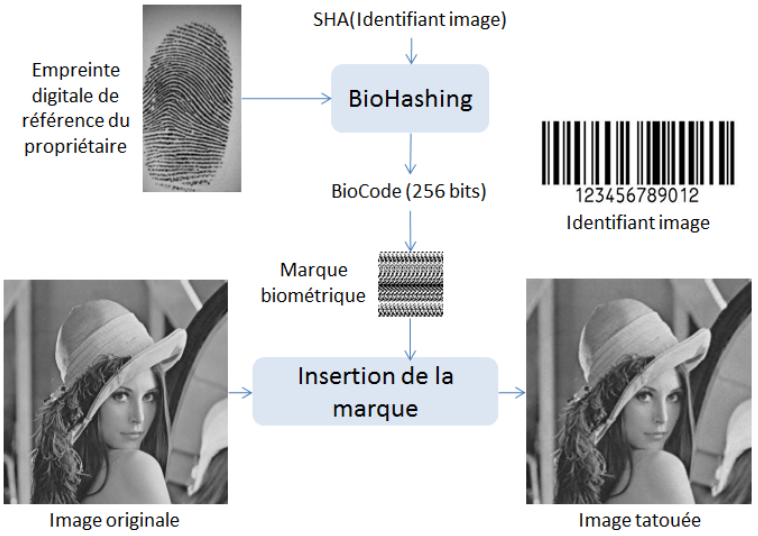

Figure 3: Insertion de la marque biométrique

3.4 Vérification de la marque biométrique

Afin de vérifier le propriétaire d'une image, l'autorité détermine l'identifiant de l'image concernée (par comparaison de l'image présentée dans sa base de données par des techniques d'indexation d'images). Le propriétaire présumé fournit sa donnée biométrique (par le biais d'une communication chiffrée) et l'autorité calcule le BioCode capturé. Le BioCode extrait de l'image est obtenu par vote majoritaire des bits de la marque biométrique de l'image tatouée (pour être robuste à d'éventuelles altérations de celle-ci). La comparaison des deux BioCodes avec la distance de Hamming permet de décider si l'utilisateur est bien le propriétaire de l'image.

3.5 Analyse de performance

Dans cette section, nous analysons dans un premier temps la robustesse du schéma de tatouage vis à vis d'attaques. Dans un second temps, nous analysons la performance de la reconnaissance biométrique malgré différentes attaques de l'image tatouée. Dans la section suivante, nous présentons le protocole expérimental mis en en place dans cette étude.

3.5.1 Protocole expérimental

Nous avons utilisé une base de données d'empreintes digitales issue de la compétition FVC2002. Elle est composée de 800 images provenant de 100 personnes avec 8 échantillons d'empreintes digitales de chaque utilisateur. Il s'agit de la base de données FVC2002 DB2 avec une résolution de l'image est 296×560 pixels acquise avec un capteur optique "FX2000" par Biometrika. Comme FingerCode, nous avons utilisé le modèle de Gabor

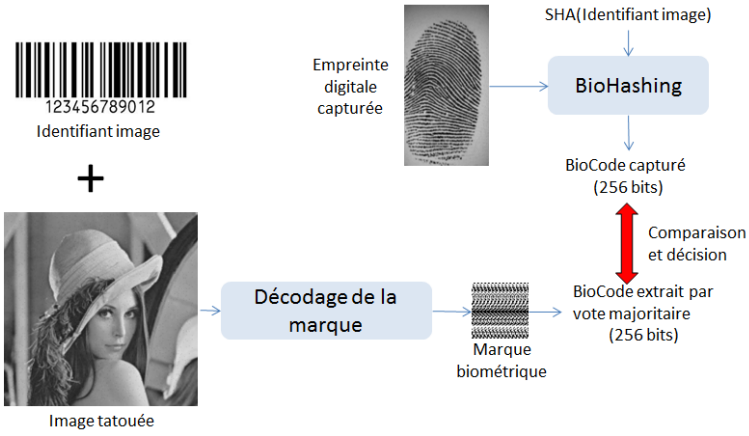

Figure 4: Décodage de la marque biométrique

[MM96] de taille $n = 512$ (16 niveaux et 16 orientations). Ces caractéristiques sont très bien connues et permettent une bonne analyse de la texture d'une empreinte digitale. Pour chaque utilisateur, nous avons utilisé le premier échantillon de FingerCode comme modèle de référence de l'utilisateur. Les échantillons restants de l'utilisateur et ceux des autres utilisateurs sont utilisés pour tester le schéma proposé. Les BioCodes sont de taille $m = 256$ bits. Afin de quantifier la performance de notre approche, nous avons calculé 1400 comparaisons (avec la distance de Hamming) entre le BioCode de référence et le BioCode capturé (des utilisateurs légitimes et imposteurs).

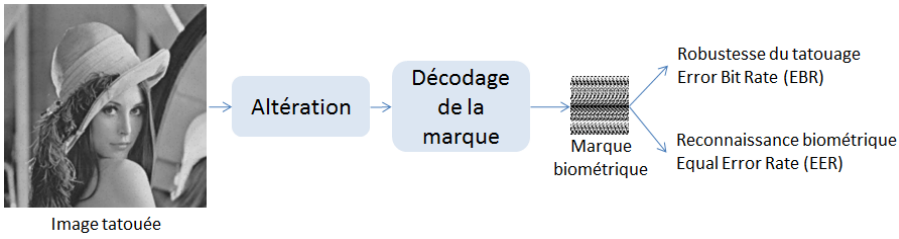

Figure 5: Procédure d'évaluation de l'approche proposée

Le protocole d'évaluation est détaillé dans la figure 5. Nous appliquons plusieurs altérations sur l'image tatouée (illustrées dans la Figure 6) simulant des modifications volontaires ou involontaires. La robustesse de l'algorithme de tatouage est estimée à partir de la métrique EBR (Error Bit Rate) définie ci-dessous. La performance de la reconnaissance biométrique est évaluée par l'EER. Ce taux calcule le ratio où les erreurs entre les utilisateurs légitimes rejetés à tort et les imposteurs acceptés à tort sont égaux.

$$EBR = \frac{\sum\sum C(x,y) \oplus \tilde{C}(x,y)}{M.N}$$

où $C(x, y)$ est la valeur de la marque initiale au pixel (x, y), $\tilde{C}$ est la marque décodée, M et N sont respectivement le nombre de lignes et colonnes de la marque (dans notre étude, N=M=64), le symbole $\oplus$ correspond à l'opérateur XOR.

Figure 6: Alterations de l'image tatouée

3.5.2 *Résultats expérimentaux*

Dans un premier temps, nous étudions la robustesse de la méthode de tatouage d'images face à différentes altérations illustrées dans la figure 6. La figure 7 présente l'évolution de l'EBR en fonction de différentes altérations. Nous pouvons constater que le schéma de tatouage est complètement invariant à l'altération par contraste. Différentes altérations engendrent une faible modification de la marque comme la luminance, le rognage et le tampon. Les autres altérations engendrent une modification importante de la marque.

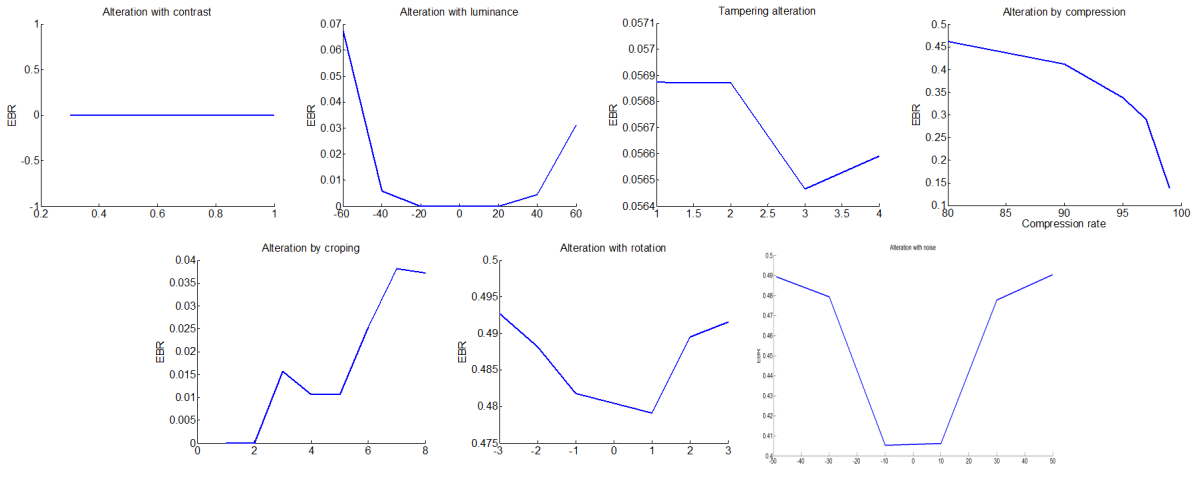

Figure 7: Évolution de l'EBR pour les différentes altérations

La figure 8 donne le taux de reconnaissance biométrique du propriétaire en présence d'attaque en considérant la valeur de l'EER. Le taux d'erreur est nul dans la plupart des cas sauf en cas de rotation de l'image ou un ajout très conséquent de bruit. Ceci montre la bonne robustesse de la méthode proposée.

No alteration	Crop 1 line	Crop 2 lines	Crop 3 lines	Crop 4 lines
0%	0%	0%	0%	0%
Crop 5 lines	Crop 8 lines	Noise -60	Noise -40	Noise -20
0%	0%	39.2%	30.3%	0.2%
Noise +20	Noise +40	Noise +60	Contrast 1	Contrast 0.7
0%	30.8%	40.5%	0%	0%
Contrast 0.5	Contrast 0.3	Luminance -60	Luminance -40	Luminance -20
0%	0%	0%	0%	0%
Luminance +20	Luminance +40	Luminance +60	Compression 99%	Compression 97%
0%	0%	0%	0%	0%
Compression 95%	Compression 90%	Compression 80%	Tampering	Rotation -3°
0%	0.2%	11%	0%	48.5%
Rotation -2°	Rotation -1°	Rotation +1°	Rotation +2°	Rotation +3°
48%	42.1%	29%	47.6%	47.1%

Figure 8: Performance de la reconnaissance biométrique

3.6 Analyse de sécurité et protection de la vie privée

3.6.1 Modèle de l'attaquant

On propose de se mettre dans le modèle de l'attaquant le plus fort qui soit : attaquant actif. La confidentialité des communications, l'authentification des interlocuteurs et l'intégrité des données sont assurées par les protocoles standards de cryptographie (canal SSL). Sans perte de généralité, on peut supposer que l'attaquant est donc passif durant les télécommunications et actif pour ses tentatives de création de documents légitimes, d'usurpation d'identité et de traçabilité des achats d'un client.

3.6.2 Analyse

Le client peut à tout moment, en fournissant son empreinte digitale et l'image en sa possession, vérifier qu'il est bien le propriétaire d'une image. Malgré des altérations de celle-ci, les résultats de reconnaissance biométrique dans la Section 3.5 ont été très satisfaisants. On en conclut que l'exigence E_1 est respectée. Compte tenu des propriétés de l'algorithme de BioHashing, le BioCode est différent pour chaque média, il est donc impossible de pouvoir retrouver l'empreinte digitale du client ou de relier différents média

au même client ; l'exigence E_2 se trouve donc satisfaite. Le FingerCode d'un client n'est divulgué qu'à l'autorité, nous considérons cet aspect raisonnable puisqu'elle est considérée ici comme un tiers de confiance. On peut imaginer d'autres variantes afin que l'autorité de confiance n'ait pas accès au FingerCode, mais cela présuppose que le client soit capable de calculer lui-même son BioCode, ce qui peut être également discutable. L'exigence E_3 est partiellement respectée, car seule l'autorité de confiance a accès aux informations personnelles du client. La vérification ne peut-être réalisée qu'avec l'empreinte digitale du client, l'exigence E_4 est donc respectée. Dans le schéma proposé, la marque biométrique est calculée à partir du FingerCode fourni par l'utilisateur et l'identifiant connu seul par l'autorité. Le client est en incapacité de générer une autre marque biométrique valable pour une autre image. L'exigence E_5 est donc respectée. Finalement, la vérification se fait par l'autorité à partir du FingerCode capturé fourni par le client et l'identifiant stocké par l'autorité. Le fournisseur ne joue aucun rôle dans la vérification, il ne peut donc pas nier avoir vendu ce média. Ainsi, l'exigence E_6 est totalement respectée.

4 Conclusion et perspectives

Dans ce papier, nous avons proposé une nouvelle méthode permettant de réaliser une preuve de propriété d'un média en utilisant une vérification biométrique respectueuse de la vie privée. Plusieurs altérations simulant une attaque volontaire ou involontaire de l'image tatouée ont montré la bonne robustesse du schéma proposé pour la vérification biométrique. Les perspectives de cette étude concernent l'utilisation de schémas de tatouage d'images plus robustes ainsi que le remplacement du code à répétition par d'autre familles de codes correcteurs plus complexes.

Références

[BF04] Paul Blythe and Jessica Fridrich. Secure digital camera. In *in Proceedings of Digital Forensic Research Workshop (DFRWS)*, pages 17–19, 2004.

[Faz06] Nelly Fazio. *On Cryptographic Techniques for Digital Rights Management.* PhD thesis, New York University, September 2006.

[KD07] Nikos Komninos and Tassos Dimitriou. Protecting biometric templates with image watermarking techniques. In Seong-Whan Lee and StanZ. Li, editors, *Advances in Biometrics*, volume 4642 of *Lecture Notes in Computer Science*, pages 114–123. Springer Berlin Heidelberg, 2007.

[MM96] B. S. Manjunath and W.Y. Ma. Texture features for browsing and retrieval of image data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18 :37–42, 1996.

[Moh09] Saraju P. Mohanty. A secure digital camera architecture for integrated real-time digital rights management. *Journal of Systems Architecture*, 55(10-12) :468 – 480, 2009.

[TNG04] A.B.J. Teoh, D. Ngo, and A. Goh. Biohashing : two factor authentication featuring fingerprint data and tokenised random number. *Pattern recognition*, 40, 2004.

[VSN+06] M. Vatsa, R. Singh, A. Noore, M. Houck, , K. Morris, Mayank Vatsa, Richa Singh, Afzel Noore, and Keith Morris. Robust biometric image watermarking for fingerprint and face template protection, 2006.

[WS11] Zhang Wenyin and Frank Y. Shih. Semi-fragile spatial watermarking based on local binary pattern operators. *Elsevier journal on Optics Communications*, pages 3904–3912, 2011.