Un panorama des méthodes de tatouage permettant d'assurer un service d'intégrité pour les images[1]

An overview of watermarking algorithms for image authentication

par Christian REY, Jean-Luc DUGELAY

Institut EURECOM, Dept. of Multimedia Communications 2229, route des Crêtes B.P. 193, 06904 Sophia Antipolis Cedex.
Tél : 33 04 93 00 26 41 Fax : 33 04 93 00 26 27 email. jean-luc.dugelay@eurecom.fr URL. http://www.eurecom.fr/~image

résumé et mots clés

Le « watermarking » ou tatouage d'image a connu, ces dernières années, un essor spectaculaire. Initialement développé pour renforcer la protection des droits d'auteur des documents multimédia, il tend de plus en plus a être utilisé pour remplir d'autres fonctions de sécurité, notamment des fonctions d'intégrité, ou des services d'information. L'objectif de cet article est de dresser un panorama des différentes méthodes permettant d'assurer un service d'intégrité adapté aux images. Contrairement aux techniques classiquement employées en sécurité pour assurer cette fonction, la plupart des méthodes proposées par la communauté « watermarking » privilégient une intégrité en termes de contenu à une intégrité stricte. Nous introduisons dans ce papier cette notion d'intégrité sémantique particulière aux images, ainsi que les critères à prendre en considération pour construire un système d'authentification performant. Plusieurs algorithmes significatifs sont détaillés afin de présenter les notions de base fréquemment usitées.

Traitement d'image, sécurité, cryptographie, tatouage d'image, intégrité, article de synthèse, état de l'art.

abstract and key words

Watermarking has become a popular technique for copyright enforcement and image authentication. The aim of this paper is to present an overview of emerging techniques for image tamper detection. Compared to the techniques and protocols usually employed for security to perform this task, most of the proposed methods based on watermarking place a particular emphasis on the notion of content authentication rather than strict integrity. In this paper, we introduce the notion of image content authentication and the features required to design an effective authentication scheme. We present some algorithms, and introduce frequently used key techniques.

Image processing, security, cryptography, watermarking, content authentication, review, state of the art.

1. Les activités de recherche en intégrité d'image de l'Institut Eurécom sont en partie soutenues par le projet RNRT Aquamars (http:www.telecom.gouv.fr/rnrt/wprojets.htm)[3]

1. introduction

1.1. généralités sur le tatouage numérique

La révolution numérique, l'explosion des réseaux de communication et l'engouement sans cesse grandissant du grand public pour les nouvelles technologies de l'information entraînent une circulation accrue des documents multimédia (images, vidéos, textes, sons, *etc*.). L'ampleur de ce phénomène est telle que des questions essentielles se posent désormais quant à la protection et au contrôle des données échangées. En effet, de par leur nature numérique, les documents multimédia peuvent être dupliqués, modifiés, transformés et diffusés très facilement. Dans ces conditions, il devient donc nécessaire de mettre en œuvre des systèmes permettant de faire respecter les droits d'auteur, de contrôler les copies et de protéger l'intégrité des documents. Dans ce contexte, le « watermarking » [7], [16] (ou tatouage numérique) est très rapidement apparu comme la solution « alternative » pour renforcer la sécurité des documents multimédia.

L'idée de base du « watermarking » est de cacher dans un document numérique une information subliminale (*i.e.* invisible ou inaudible suivant la nature du document) permettant d'assurer un service de sécurité (copyright, intégrité, traçabilité, non répudiation, *etc*.) ou à but d'information. Une des particularités du tatouage numérique par rapport à d'autres techniques, comme par exemple un stockage simple de l'information dans l'en-tête du fichier, est que la marque est liée de manière intime et résistante aux données. De ce fait, le tatouage est théoriquement indépendant du format de fichier et il peut être détecté ou extrait même si le document a subi des modifications ou s'il est incomplet.

Le problème de l'intégrité est encore peu abordé par la communauté « watermarking » et de nombreuses questions restent ouvertes. On peut par exemple se demander s'il est préférable d'avoir recours à un tatouage fragile plutôt qu'à un tatouage robuste, ou bien encore opter pour une toute autre solution ? D'autre part, un service d'intégrité remet partiellement en cause certains paramétrages communément établis en tatouage d'image pour assurer une fonction plus classique de sécurité de type « droits d'auteur », notamment en termes de quantité et nature des informations cachées (pour le copyright, la marque est indépendante de l'image et est usuellement un identifiant codé sur 64 bits), ainsi qu'en terme de robustesse.

1.2. notions d'intégrité

La notion d'intégrité est un concept bien connu en sécurité. Sa définition repose sur une décision binaire qui garantit que les données reçues sont rigoureusement identiques à celles émises. Cette définition est applicable à tout type de documents numériques, néanmoins, dans la pratique elle s'avère être beaucoup trop stricte et inadaptée pour les documents multimédia. En effet, l'interprétation que l'on a d'une image dépend principalement des éléments la constituant plutôt que des valeurs numériques des pixels ou de sa résolution. En d'autres termes, le problème de l'intégrité des images se pose principalement en termes de contenu sémantique ; c'est-à-dire la détection des modifications du document pouvant engendrer une gêne dans sa visualisation et/ou une erreur dans son interprétation (modification de la légende, disparition d'un visage, *etc*.). Dans le but d'assurer un service d'intégrité approprié aux images, il est donc primordial de distinguer les manipulations malveillantes consistant à détourner le contenu initial de l'image, des manipulations liées à son utilisation ou son stockage sous une forme numérique (conversion de format, compression, ré-échantillonage, filtrage, *etc*.) réalisés par des fournisseurs de contenu ou les utilisateurs eux-mêmes. Malheureusement cette distinction n'est pas toujours aisée d'un point de vue informatique et dépend en partie du type d'image et de son utilisation. Par exemple, dans le cas particulier de l'imagerie médicale, des manipulations anodines, comme une simple compression, voire le processus de tatouage lui-même, peuvent causer la disparition de certains signes visibles d'une pathologie faussant alors le diagnostic du médecin. Dans ce contexte, l'utilisation de méthodes dites classiques sera plus approprié pour garantir une intégrité stricte du document.

1.3. exemples classiques de manipulations malveillantes

Dans notre société, les messages véhiculés par les images ont un impact considérable. En effet, le réalisme d'une photographie est tel que nous avons tendance à prendre pour réelles des scènes qui ne le sont pas (toutes les images, y compris celles réalisées en toute innocence, ont la capacité d'être détournées de leur sens). Les manipulations, qui avant, nécessitaient des moyens coûteux sont désormais à la portée de tout le monde et les progrès de la technique et du tout numérique les rendent quasi indécelables.

Dans ce contexte, un service d'intégrité d'image n'a bien évidemment pas la prétention de vérifier la véracité des événements, mais de déceler des manipulations qui auraient pu y être apportées *a posteriori* (*i.e.* entre la prise de la photographie et sa diffusion) dans le but de détourner le contenu de l'image ou de rendre impossible toute interprétation. Nous donnons ci-après quelques exemples célèbres de manipulations intentionnelles d'images (sources : Ça m'intéresse – septembre 2000 [6]). La couverture du magazine « Time » du 27 juin 1994 est un bel exemple de falsification d'image (figure 1), ainsi qu'un véritable scandale journalistique qui a fait couler beaucoup d'encre à l'époque. Les éléments ajoutés à la photographie originale

Figure 1 – Exemple de falsification d'image (l'affaire O.J. Simpson).

Image originale extraite du film « La guerre des étoiles » **Image retouchée**

Figure 2 – Exemple de manipulation modifiant le contenu de l'image.

d'O.J. Simpson (flou, noircissement du visage, halo obscur), créés de toutes pièces au moyen d'un logiciel de retouche à des fins de manipulation, exposèrent au grand jour le problème de la numérisation des images. Un autre exemple qui a fait le tour du monde est celui du recadrage de la place Tien An Men. En sélectionnant le personnage et le premier char, le recadrage de cette photographie n'en modifie pas la signification générale, mais en dramatise et en mystifie l'action. Une autre affaire de photographie truquée célèbre est celle diffusée en 1995, dans l'émission de France 3, « La marche du siècle », où de jeunes « beurs » avaient été transformés à leur insu en redoutables intégristes (*source :* Le Monde Diplomatique [37]). Plus récemment, une photographie publiée en une du quotidien autrichien « *Neue Kronen Zeitung* », prétend illustrer l'agressivité des manifestants opposés à l'entrée du parti de Haider dans le gouvernement autrichien. Par un truquage numérique on a recadré la photographie et raccourci la distance entre un manifestant et un policier apparemment directement frappé. En réalité, comme l'atteste l'image originale diffusée par l'agence Reuters [35], une distance de près de deux mètres séparait les deux protagonistes. Aussi l'utilisation comme élément à charge par l'image, l'audio ou la vidéo devient plus que douteuse et critiquable à l'heure où les caméras de surveillance envahissent les villes, les stades et les routes.

1.3.1. schéma générique d'un système d'authentification d'image

On se propose de définir un schéma générique d'un système d'authentification d'image (dont différentes formulations ont été initialement proposées par Wu et Liu [47] et Lin et Chang [25]). Pour être efficace, ce dernier doit satisfaire les critères suivants :

– **sensibilité** : le système doit être capable de déceler des manipulations pouvant modifier l'interprétation que l'on a d'une image, telles que des recadrages (crop) ou des retouches locales (exemple figure 2) ;

– **tolérance** : le système doit être tolérant vis-à-vis des algorithmes de compression avec pertes tels que Jpeg, et plus généralement vis-à-vis des manipulations bienveillantes (générées, par exemple, par les fournisseurs de contenu multimédia ou un utilisateur de bonne foi) ;

– **localisation des régions altérées** : le système doit être en mesure de donner à l'utilisateur une information visuelle permettant d'identifier rapidement les régions qui ont été manipulées ;

– **reconstruction des régions altérées** : le système doit éventuellement permettre une restauration partielle des zones de l'image qui ont été manipulées ou détruites, afin de donner à l'utilisateur la possibilité de se faire une idée sur le contenu original de ces régions.

En plus des critères précédents, d'autres contraintes techniques sont également à prendre en considération :

– **mode de stockage** : il est préférable de cacher les données d'authentification dans l'image elle-même, sous la forme d'un tatouage, plutôt que dans un fichier séparé comme dans le cas d'une signature externe ;

– **mode d'extraction** : suivant que les données d'authentification sont dépendantes ou non de l'image, on optera pour un mode d'extraction du tatouage aveugle ou semi-aveugle [7]. En mode d'extraction aveugle, la marque représentant les données d'authentification est récupérée à partir de l'image marquée seule (éventuellement manipulée), alors qu'en semi-aveugle il s'agit principalement de vérifier la présence de telle marque dans une image (*via* un score de corrélation). Il est bien évident qu'un mode d'extraction non aveugle est dénué de sens pour un service d'intégrité dans la mesure où il fait appel à l'image originale ;

– **algorithme asymétrique** : Contrairement aux services de sécurité plus classiques comme le « copyright » où l'on peut se contenter d'une même clé (privée) pour l'insertion et l'extraction de la marque, un service d'intégrité nécessite de préférence l'utilisation d'un algorithme de tatouage asymétrique (ou de chiffrement, selon le cas) dans la mesure où tout un chacun doit pouvoir s'assurer de l'intégrité d'une image ;

– **visibilité** : les données d'authentification doivent être invisibles (dans les conditions normales de visualisation). Il s'agit de faire en sorte que l'impact visuel du marquage (*i.e.* distorsion) soit le plus faible possible afin que le document marqué reste fidèle à l'original ;

– **robustesse et sécurité** : les données d'authentification doivent être protégées par des méthodes de chiffrement de manière à éviter qu'elles soient falsifiées ou manipulées ;

– **protocoles** : enfin, les protocoles tiennent également une place prépondérante dans tout système d'authentification d'image. En effet, l'algorithme ne permet pas à lui seul de garantir l'authenticité d'une image. Il est nécessaire de définir en plus un ensemble de spécifications décrivant les conventions et les règles du système, comme par exemple la gestion des clés ou bien encore éviter qu'une image déjà protégée, puisse l'être à nouveau, a fortiori si elle a été manipulée.

2. état de l'art

2.1. introduction

Cette section n'a pas pour objectif de dresser un panorama complet et exhaustif des différentes techniques permettant d'assurer un service d'intégrité pour les images. Néanmoins, le but de cette partie est de présenter dans les grandes lignes plusieurs méthodes significatives du domaine afin d'introduire progressivement les notions clés associées à ce type de service.

Les systèmes d'authentification des images peuvent être regroupés de plusieurs manières suivant qu'ils assurent un service d'intégrité stricte ou bien une intégrité en termes de contenu, suivant le mode de stockage des données d'authentification (*i.e.* tatouage ou signature externe), ou bien encore selon la nature des informations qu'ils enfouissent dans le document à protéger.

2.2. tatouages fragiles

2.2.1. principe

Les premières méthodes proposées pour assurer un service d'intégrité étaient basées sur l'utilisation d'un tatouage fragile, par

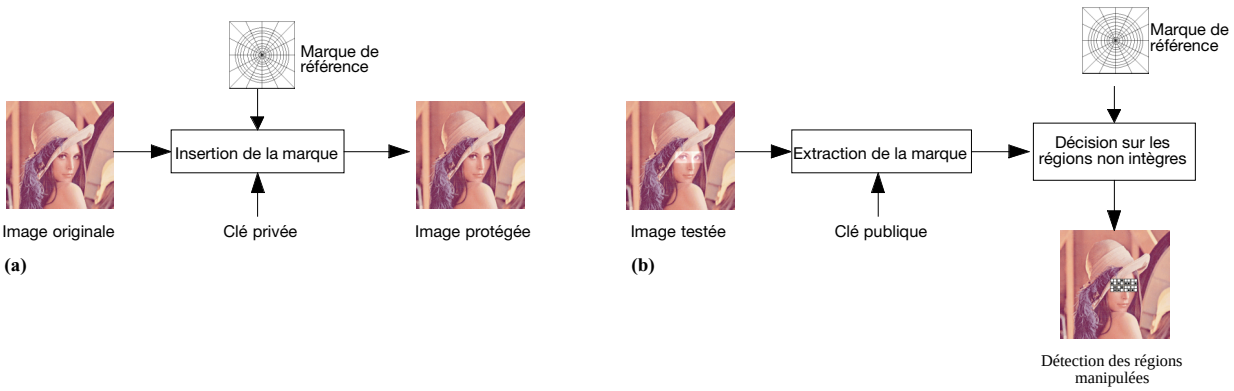

Figure 3 – Schéma général d'un système d'authentification basé sur un tatouage fragile a) Protection de l'image, b) Authentification de l'image.

opposition au tatouage robuste classiquement utilisé pour la protection des droits d'auteur. Le principe de ces approches est d'insérer une marque ou un logo binaire (généralement prédéfini et indépendant des données à protéger [48]) dans l'image d'origine de telle manière que les moindres modifications apportées à l'image se répercutent également sur la marque insérée (figure 3.a). Pour vérifier l'intégrité d'une image, il suffit alors de vérifier localement la présence de cette marque (figure 3.b).

2.2.2. insertion de « checksums » dans les LSB

Une des premières techniques utilisées pour vérifier l'intégrité d'une image visait à insérer des valeurs de « checksums » dans les bits les moins significatifs (LSB) des pixels de l'image. L'algorithme proposé par Walton [42] en 1995 consiste à sélectionner, de manière pseudo-aléatoire (en fonction d'une clé), des groupes de pixels et de calculer, pour chacun d'eux, une valeur de « checksum ». Ces valeurs sont obtenues à partir des nombres formés par les 7 bits les plus significatifs (MSB) des pixels sélectionnés, et sont ensuite insérées sous forme binaire au niveau des bits de poids faible. Ci-après, de manière plus détaillée l'algorithme tel qu'il était proposé à l'origine :

Algorithme 1 Etape d'insertion

1. Soit N suffisamment grand ;
2. Diviser l'image en blocs de taille 8×8 pixels ;
3. Pour chaque bloc B_i :

– définir un ordre de parcours pseudo-aléatoire (selon par exemple une clé secrète et l'indice du bloc B_i) des 64 pixels (p_1, $p_2, \ldots, p_{64}$) ;

– générer une séquence pseudo-aléatoire de 64 entiers (a_1, a_2, $\ldots, a_{64}$) du même ordre de grandeur que N ;

– la valeur de checksum S est alors calculée de la manière suivante :

$$S = \sum_{j=1}^{64} (a_j.g(p_j)) \bmod N$$

avec $g(p_j)$ le niveau de gris du pixel p_j en ne tenant compte que des 7 MSB.

– coder et crypter S en binaire ;

– insérer la séquence binaire résultante au niveau des LSB des pixels du bloc

L'algorithme de vérification est dual de celui d'insertion. Il consiste à vérifier pour chaque bloc, la valeur de « checksum » recalculée à partir des MSB des pixels de l'image testée, avec celle de l'image originale codée au niveau des LSB.

Cette méthode garantit d'une part, en insérant les données d'authentification directement au niveaux des LSB de l'image, une distorsion visuelle minime, quasi imperceptible par l'œil humain. D'autre part, elle a l'avantage d'être simple, rapide et sensible à la moindre modification de l'image (*i.e.* réponse binaire équivalente à une intégrité stricte). Si on échange, par exemple, les MSB de deux pixels quelconques d'un même bloc, la valeur de S s'en trouvera automatiquement modifiée car chaque pixel p_j est multiplié par un coefficient a_j différent. De plus, l'ordre de parcours des pixels p_j ainsi que les valeurs des coefficients a_j sont dépendants du bloc, ce qui rend impossible un éventuel « copier/coller » entre deux blocs différents d'une même image.

Le lecteur attentif remarquera cependant, qu'ainsi défini, il est possible avec cette méthode d'intervertir deux blocs homologues (*i.e.* de même position) de deux images protégées avec la même clé, sans que le système ne décèle une perte d'intégrité (depuis différentes améliorations ont été proposées [11] pour pallier à ce type d'attaque). Par contre si l'image est légèrement recadrée ou compressée, le système détecte une perte d'intégrité alors que le contenu sémantique de l'image reste inchangé.

2.2.3. « Self-embedding »

Fridrich et Goljan [12] ont, quant à eux, développé une technique utilisant également les LSB comme support, mais dans le but, cette fois-ci, de cacher suffisamment d'informations afin de pouvoir non seulement déceler d'éventuelles manipulations, et aussi de permettre une reconstruction partielle des régions détériorées. L'idée de base consiste à découper l'image en blocs 8×8, à en calculer les coefficients DCT (Transformée en Cosinus Discrète) [33] en ne tenant compte bien évidemment que des MSB. Ces coefficients DCT sont ensuite quantifiés à l'aide de la table de quantification correspondant à une compression Jpeg d'une qualité de l'ordre de 50 %. La matrice quantifiée résultante est alors encodée sur 64 bits et insérée au niveau des LSB des pixels d'un autre bloc. Le bloc servant de support au tatouage doit être suffisamment éloigné afin d'éviter qu'une manipulation locale de l'image ne détériore à la fois l'image et les données de reconstruction correspondantes.

Comme pour toutes les méthodes de tatouage utilisant les LSB comme support, l'impact visuel est très faible. Par contre, la qualité des régions restaurées est nettement inférieure à celle d'une compression Jpeg 50 %, mais largement suffisante pour informer l'utilisateur sur le contenu original de ces régions. Les auteurs ont également proposé une variante afin d'améliorer légèrement la qualité de la reconstruction en utilisant cette fois-ci les deux bits de poids faible comme support (la matrice quantifiée étant alors codée sur 128 bits). La reconstruction est certes meilleure, mais l'image tatouée perd sensiblement en qualité.

Le principal inconvénient de cette méthode est lié à la nature très fragile du tatouage qui ne garantit pas, dès lors que plusieurs régions de l'image ont été manipulées, une restauration correcte. En effet, les données de reconstruction correspondant à un bloc erroné peuvent elles aussi être altérées si les LSB les supportant

ont eux aussi été modifiés. Ce problème est d'autant plus vrai lorsque l'image subie des manipulations globales, même « faibles », comme un filtrage passe-bas ou une compression Jpeg. D'une manière générale, on peut donc légitimement se poser la question de l'intérêt des méthodes de tatouages fragiles vis à vis des techniques cryptographiques classiques, dans la mesure où elles ne garantissent également qu'une intégrité stricte.

3.3. tatouages semi-fragiles

Face à ce constat de semi-échec, les recherches s'orientent actuellement vers des approches dites semi-fragiles. Les méthodes ayant recours à un tatouage semi-fragile se distinguent des méthodes fragiles dans la mesure où elles offrent une robustesse accrue face à certaines manipulations d'image. L'objectif recherché est de pouvoir discriminer des opérations malveillantes, comme par exemple l'ajout ou la suppression d'un élément important de l'image, de transformations globales « raisonnables » ne portant pas atteinte au contenu sémantique de l'image.

L'utilisation de telles méthodes est principalement motivé par le fait que les images sont généralement transmises et stockées sous une forme compressée et que pour la majorité des applications, les pertes liées au processus de compression n'affectent pas l'intégrité de l'image au sens de son interprétation.

3.3.1. exemple de méthode transparente à la compression Jpeg

Dans [25], Lin et Chang proposent un algorithme d'authentification robuste à la compression Jpeg. Les auteurs ont mis en évidence et démontré deux propriétés d'invariance des coefficients DCT vis-à-vis de la compression Jpeg [30].

La première propriété énonce que si on donne à un coefficient DCT, quel qu'il soit, une valeur entière multiple d'un pas de quantification prédéfini Q'_m supérieur à tous les pas de quantification possibles d'une compression Jpeg acceptable (*i.e.* facteur qualité de 50 % environ), alors cette valeur peut être recalculée exactement après une compression Jpeg acceptable. La deuxième propriété définit une règle d'invariance de la relation d'ordre entre les coefficients homologues de deux blocs DCT vis-à-vis de la compression Jpeg. En effet, lors de la compression, les différents blocs DCT d'une image sont tous divisés par la même table de quantification, de ce fait la relation qui lie les coefficients de mêmes coordonnées de deux blocs reste inchangée après le processus de quantification. La seule exception est que dans certains cas, des inégalités strictes peuvent devenir de simples égalités, par le biais de la quantification.

Le système d'authentification proposé par Lin et Chang repose donc sur ces deux propriétés. La première est utilisée pour définir un support de tatouage robuste à la compression Jpeg, tandis que la seconde sert à générer les données d'authentification proprement dites. Les étapes d'insertion et d'authentification peuvent se résumer ainsi :

Algorithme 2.a Génération des bits d'authentification

1. Découper l'image originale en blocs 8×8
2. Apparier les blocs deux par deux en fonction d'une clé secrète
3. Pour chaque paire de blocs (p, q) :

– sélectionner un ensemble B de n coefficients DCT (autres que la composante continue) ;

– générer la signature binaire ϕ de la paire de blocs à l'aide de la règle suivante :

$$\phi(\nu) = \begin{cases} 1, & F_p(\nu - F_q(\nu) \geqslant 0 \\ 0, & F_p(\nu) - F_q(\nu) < 0 \end{cases}$$

avec $\nu \in B$ et $F(\nu)$ la valeur du coefficient ν.

– insérer les bits d'authentification suivant l'algorithme 2.b.

La signature binaire obtenue est ensuite cachée en partie dans chacun des deux blocs de la paire. L'algorithme de tatouage utilisé est relativement simple puisqu'il s'agit de définir une relation d'égalité entre les LSB des coefficients DCT prédéfinis avec les bits de la signature.

Algorithme 2.b Insertion du tatouage

1. Sélectionner un ensemble E, de $\dfrac{n}{2}$ coefficients DCT, avec $E \cap B = \emptyset$;
2. Pour cacher un bit d'authentification $\phi(\nu)$ dans un coefficient DCT ω :

Soit $f'_p(\omega) = \left[\dfrac{F_p(\omega)}{Q'_m(\omega)}\right]$

$$\widetilde{F}_p(\omega) = \begin{cases} f'_p(\omega) \cdot Q'_m(\omega), \text{ si LSB}\left(f'_p(\omega)\right) = \phi(\nu) \\ \left(f'_p(\omega) + signe\left(\dfrac{F_p(\omega)}{Q'_m(\omega)} - f'_p(\omega)\right)\right) \cdot Q'_m(\omega), \text{ sinon.} \end{cases}$$

avec $signe(x) = 0$ si $x < 0$, 1 sinon.

La vérification de l'intégrité d'une image est réalisée simplement en extrayant les bits d'authentification des coefficients DCT recevant le tatouage et en comparant la signature extraite avec celle obtenue à partir des blocs de l'image testée. Si les deux signatures correspondent parfaitement, la paire de blocs est alors jugée intègre, dans le cas contraire cela signifiera que l'un des deux blocs, voire les deux, ont été manipulés.

Les auteurs ont proposé de nombreuses améliorations à cette méthode, notamment l'ajout de bits de reconstruction. L'intérêt de ces bits supplémentaires est double. Ils permettent d'une part, comme leur nom l'indique, de reconstruire partiellement les blocs erronés, et d'autre part d'aider à localiser précisément les zones de l'image qui ont réellement été altérées (*i.e.* lever l'am-

biguïté sur l'identification des blocs erronés). Les bits de reconstruction sont obtenus à partir d'une version sous-échantillonnée et compressée de l'image originale, et sont ensuite insérés de la même manière que les bits d'authentification dans quatre blocs de l'image originale.

2.3.2. tatouage par région

Le tatouage par région consiste à découper l'image que l'on souhaite protéger en blocs relativement grands (de l'ordre de 64×64 pixels) et d'insérer, dans chacun d'eux, une marque « relativement robuste ». Lorsque l'on souhaite vérifier l'intégrité de l'image, on teste la présence de la marque dans les différents blocs. Dans le cas où la marque est présente avec une probabilité élevée dans chacun des blocs, on peut affirmer que l'image testée est intègre.

La technique *Variable-Watermark Two-Dimensional* (VW2D) décrite par Wolfgang et Delp [44], [45] reprend le principe décrit précédemment ; à savoir de cacher une marque binaire différente $W(b)$ dans chaque bloc b d'une image X. Ils préconisent de générer une marque binaire pseudo-aléatoire à partir de « m-sequences » [31] à la manière des travaux initiés par Shyndel *et al* [39]. L'utilisation de « m-sequences » est en effet motivée par le fait qu'elles ont d'excellentes propriétés d'auto-corrélation, ainsi qu'une très bonne robustesse à l'ajout de bruit. Dans le système d'authentification proposé, la séquence binaire $\{0, 1\}$ est transformée en une séquence de $\{-1, 1\}$, puis arrangée de manière à former un bloc de même taille que le bloc de l'image auquel elle va être modulée. La modulation de la marque et de l'image est réalisée très simplement en ajoutant ou en supprimant un niveau de gris au pixel correspondant (équation 1) :

$$Y(b) = X(b) + W(b) \tag{1}$$

avec X l'image originale, et Y l'image tatouée.

Ensuite, pour déterminer si la marque recherchée est bien présente dans un bloc, on calcule un score statistique δ (équation 3) basé sur un calcul de corrélation (équation 2) entre l'image (marquée et attaquée) et la marque :

$$A(b) \cdot B(b) = \sum_i \sum_j A(i,j)B(i,j) \tag{2}$$

$$\delta(b) = Y(b) \cdot W(b) - Z(b) \cdot W(b) \tag{3}$$

avec Z l'image à tester (lors de la vérification) la marque W est supposée connue.

Si $\delta < T$, avec T un seuil fixé par l'utilisateur, le bloc est alors jugé intègre. En jouant sur la valeur de T, on tolère des changements plus ou moins importants dans l'image. De ce fait il est possible d'affiner la détection en définissant plusieurs seuils correspondant à plusieurs niveaux de dégradation pour les blocs (par exemple : intègre, légèrement altéré, très dégradé, complètement modifié).

Cependant, dans la pratique, cette méthode n'offre qu'un intérêt limité dans la mesure où il est nécessaire de stocker au minimum, pour chaque bloc b d'une image, le résultat de la corrélation entre le bloc tatoué $Y(b)$ et la marque cachée $W(b)$.

Fridrich [13], [14] propose une technique similaire, mais préconise, pour des raisons de sécurité, de rendre le tatouage dépendant de la région de l'image dans laquelle il est inséré. La marque binaire utilisée correspond à un signal pseudo-aléatoire généré à partir d'une clé secrète, du numéro du bloc et d'un M-tuplet de bits représentatifs de la portion d'image considérée. Chaque bloc est ensuite tatoué en utilisant une technique d'étalement de spectre, similaire à celle proposée par Ó Ruanaidh [38]. D'après l'auteur, la marque offre une bonne robustesse aux opérations classiques de traitement d'image telles que de petits ajustements de contraste ou de luminosité, l'ajout de bruit, l'application de filtres passe-bas ou passe-haut, l'égalisation d'histogramme, ou bien encore une compression Jpeg de l'ordre de 50 %, permettant ainsi de distinguer des changements liés à l'utilisation d'une image, de manipulations malveillantes.

2.3.3. tatouage de caractéristiques de l'image

L'idée de base de cette méthode [36] consiste à extraire certaines caractéristiques de l'image originale et à les cacher ensuite dans l'image sous la forme d'un tatouage robuste et invisible au sens classique du « copyright » [8], [9]. Lorsque l'on souhaite vérifier l'intégrité d'une image, on compare simplement les caractéristiques de cette image avec celles de l'image originale contenues dans le tatouage. Si les caractéristiques sont identiques, cela signifiera que l'image n'a pas été manipulée, sinon les différences indiqueront les régions qui ont été altérées (figure 4).

Le choix des caractéristiques de l'image est primordial dans la mesure où il va conditionner les manipulations que l'on pourra détecter et celles qu'on laissera passer. De plus, ce choix dépend également du type d'image considéré (peinture, image satellite, image médicale, photo, *etc.*), ainsi que de l'application visée. D'une manière générale, on sélectionne les traits de l'image en fonction de leur stabilité face aux différentes attaques. Typiquement, on recherchera des caractéristiques qui sont invariantes face à une compression Jpeg, mais sensibles à des retouches locales de l'image, comme par exemple la luminance moyenne par bloc.

Cette technique impose également de nouvelles contraintes, principalement en termes de robustesse et de capacité d'insertion. En effet, il est impératif, d'une part, d'extraire la marque sans erreur sous peine d'avoir un taux élevé de fausses alarmes. D'autre part, la précision de la détection des régions de l'image qui ont été manipulées est directement liée à la quantité d'information cachée dans l'image. Il est donc nécessaire de trouver un

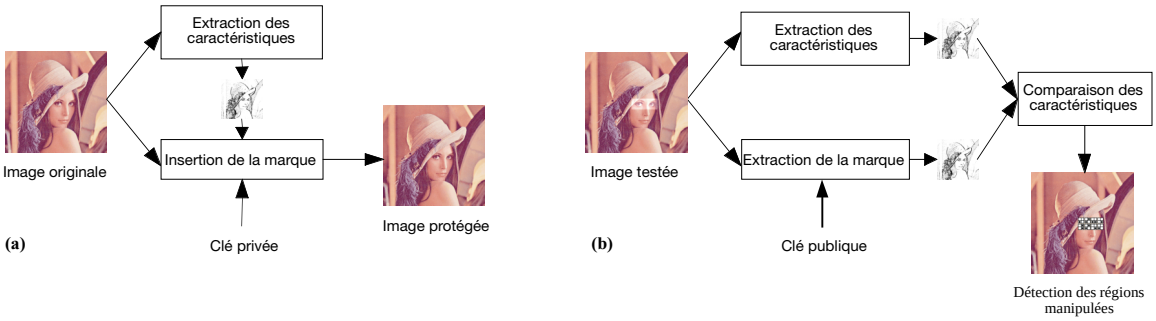

Figure 4 – Principe général d'un système d'authentification basé sur le tatouage de caractéristiques de l'image (a) Protection de l'image, (b) Vérification de l'intégrité d'une image.

bon compromis pour la taille de la marque afin de satisfaire à la fois aux deux contraintes : robustesse et sensibilité de la détection.

Une des difficultés de dissimuler les attributs caractéristiques de l'image sous la forme d'un tatouage réside dans le fait que l'image tatouée est légèrement modifiée par l'insertion de la marque elle-même. Bien que ces variations soient imperceptibles à l'œil, elles affectent légèrement les caractéristiques intrinsèques de l'image. De ce fait, les caractéristiques de l'image originale et celles de l'image tatouée ne sont plus exactement les mêmes, et on risque alors de détecter des régions altérées alors que l'image n'a pas été manipulée. Ce risque de fausses alarmes est plus ou moins important en fonction du type des caractéristiques choisi et de l'algorithme d'insertion utilisé. Ce problème a été résolu grâce à un processus de tatouage itératif [36]. Ce processus est initialisé en tatouant une première fois l'image originale avec ses propres caractéristiques. Puis, de manière itérative, on extrait les nouvelles caractéristiques de l'image tatouée que l'on insère à nouveau dans l'image originale sous la forme d'un nouveau tatouage. Seule l'image originale est marquée pour éviter d'accumuler des distorsions liées au processus de tatouage. De cette manière, grâce à ce processus itératif, les caractéristiques contenues dans le tatouage coïncident quasi parfaitement avec celles de l'image une fois tatouée.

2.3.4. autres approches

D'autres techniques sont étudiées ou en cours d'investigation. Parmi celles-ci on peut citer celle de Kundur et Hatzinakos [19], et celle de Lin et Chang [21] qui utilisent les ondelettes. Le principe de la méthode proposée par Lin et Chang est de choisir, tout d'abord, un bruit pseudo-aléatoire et une ondelette de base, qui constituent le secret du système d'authentification. Puis de décomposer l'image en 4 sous-bandes (LL, LH, HL et HH) en fonction de l'ondelette de base choisie au départ. L'étape suivante revient à substituer la sous-bande HH par le bruit pseudo-aléatoire et à effectuer ensuite la transformation en ondelettes inverse afin d'obtenir l'image tatouée. Il est intéressant de noter que le fait de modifier uniquement la sous-bande HH (*i.e.* hautes fréquences) n'entraîne pas de dégradations visibles.

Le processus d'authentification consiste alors à effectuer la même décomposition que lors de la phase d'insertion, puis à corréler la sous-bande HH obtenue avec le bruit pseudo-aléatoire. Si l'image n'a subi aucune manipulation, le résultat du test ressemblera à une matrice de points uniformément répartis. Dans le cas contraire, la distribution perdra son caractère uniforme dans les régions où l'image a été manipulée. Les auteurs font remarquer que cette méthode est « perméable » à certaines manipulations telles qu'un flou ou un rehaussement des contours, dans la mesure où les changements ne sont pas trop importants. Expérimentalement, cette méthode permet également de laisser passer une légère compression Jpeg. Par contre, les auteurs ne démontrent pas la robustesse de leur méthode face à des attaques spécifiques visant par exemple à substituer la sous-bande HH ou au contraire à la préserver (*i.e.* modifier l'image, puis réinsérer la sous-bande HH de l'image originale protégée). En d'autres termes, est-ce que le choix de l'ondelette de base comme secret est suffisant pour éviter ce type d'attaques ?

2.4. signatures externes

Les signatures externes offrent une alternative aux techniques de tatouage classiques dans le cadre d'un service de contrôle d'intégrité dans les images. Contrairement aux techniques de tatouage d'image, la marque n'est pas insérée dans l'image elle-même, mais transmise avec celle-ci sous une forme chiffrée.

On peut établir un parallèle entre l'utilisation de signatures numériques pour assurer un service d'authentification et le domaine de l'indexation d'image [29], où de nombreuses techniques ont recours à ce type d'empreintes externes (ou signatures condensées) pour retrouver des images en fonction de leur contenu. Ces signatures sont générées le plus souvent à partir d'attributs significatifs qui traduisent le contenu sémantique, tels que la couleur, la forme ou la texture.

2.4.1. fonctions de hachage

Le rôle principal des fonctions de hachage est de permettre de vérifier l'intégrité d'un document numérique sans avoir recours à l'original. Une fonction de hachage opère généralement sur un message M de longueur arbitraire pour fournir une valeur de hachage h de taille fixe. Pour qu'une telle fonction soit considérée comme sûre elle doit vérifier les propriétés suivantes :

– il est « facile » de calculer h connaissant M,

– il est « difficile » de retrouver M connaissant h,

– il est « difficile » de trouver un message M' (différent de M) ayant comme valeur de hachage $h' = h$.

En d'autres termes, une fonction de hachage sert à produire un condensé (ou une empreinte) unique, représentatif du document original. Il existe de nombreuses fonctions de hachage parmi lesquelles on peut citer : MD-4, MD-5 (Message Digest), CRC-32 (32 bits Cyclic Redundancy Check), SHA-1 (Secure Hash Algorithm) [40], *etc.*

Row-column hash function

La technique des « row-column hash function » [44] consiste à calculer une valeur de hachage pour chaque ligne et chaque colonne de l'image originale. Lorsque l'on souhaite vérifier l'intégrité d'une image, on recalcule les valeurs de hachage des lignes et des colonnes de l'image à tester et on les compare avec celles de l'image originale. Pour localiser les éventuelles disparités, il suffit d'identifier les lignes et les colonnes qui sont différentes. Cependant, dans le cas où plusieurs zones de l'image ont été modifiées, on n'est plus capable de les localiser sans ambiguïté, c'est-à-dire que des régions intègres seront considérées comme altérées (figure 5), ce qui réduit considérablement l'intérêt de cette technique.

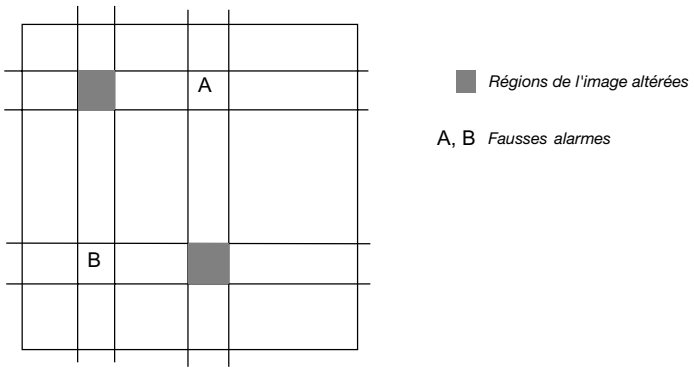

Figure 5 – Ambiguïté dans la localisation des régions altérées de l'image.

Block based hash function

Un autre algorithme utilise également des fonctions de hachage, il s'agit du Block-Based Hash function (BBH) [44]. Le principe est similaire à celui décrit précédemment, à la différence près qu'il n'opère plus sur les lignes ou les colonnes de l'image, mais sur des blocs. Ainsi lorsque l'on constate des différences dans les valeurs de hachage, il suffit de se reporter aux blocs concernés pour localiser les zones de l'image qui ont été manipulées.

Les fonctions de hachage ont la particularité d'être extrêmement sensibles à la moindre variation ; en effet il suffit de modifier la valeur d'un pixel d'un seul bit pour changer radicalement la valeur de hachage du bloc associé. Elles ne permettent donc pas de distinguer les manipulations malveillantes des manipulations bienveillantes (*i.e.* utilisateurs ou fournisseurs de contenus).

2.4.2. signature basée sur des caractéristiques de l'image

Contrairement aux techniques ayant recours à des fonctions de hachage pour générer une empreinte de l'image, certains auteurs, comme Queluz [32] ou Lin et Chang [22], [23], proposent d'extraire des caractéristiques intrinsèques de l'image, telles que les contours, et de les crypter à l'aide d'un algorithme de chiffrement asymétrique afin de les transmettre en même temps que l'image (figure 6).

Dans le cas d'un système d'authentification basé sur l'utilisation d'une signature externe, la distinction entre des manipulations innocentes et malveillantes repose principalement sur le choix des caractéristiques de l'image pour générer la signature. Ce problème est exactement le même que celui rencontré par certaines méthodes semi-fragiles, à la différence près qu'ici la contrainte de quantité d'informations, liée à la capacité de l'algorithme de tatouage, ne se pose plus (puisque les informations sont stockées dans un fichier séparé). Certains auteurs, comme Lin et Chang, suggèrent de coder la relation d'ordre entre les coefficients DCT homologues de deux blocs distincts (la méthode est la même que celle présentée au paragraphe 2.3.1). Queluz [32], quant à elle, opte pour des caractéristiques plus visuelles (principalement les contours), mais également moins stables. Pour compenser ce manque de stabilité, elle a mis en place des post-traitements complexes afin de réduire les fausses alarmes liées à une compression Jpeg.

Bhattacharjee et Kutter [5] proposent également une technique ayant recours à une signature externe ; mais plutôt que d'extraire des caractéristiques par bloc d'image, ils suggèrent de rechercher des points d'intérêts (en se basant sur les travaux de Manjunath *et al.* [26]) et de coder leurs coordonnées. La signature ainsi obtenue est ensuite chiffrée à l'aide d'un algorithme à clé privée / publique tel que RSA (méthode de chiffrement asymétrique inventée par Rivest, Shamir et Adleman [34]).

L'inconvénient majeur de ces techniques qui font appel à une signature externe pour assurer un service d'intégrité d'image est que l'image ne s'auto-suffit plus. On perd par conséquent en grande partie, l'intérêt du tatouage d'image. De plus, ces méthodes soulèvent de nouveaux problèmes comme par exemple l'authenticité de la signature, ainsi que celle du couple image/signature.

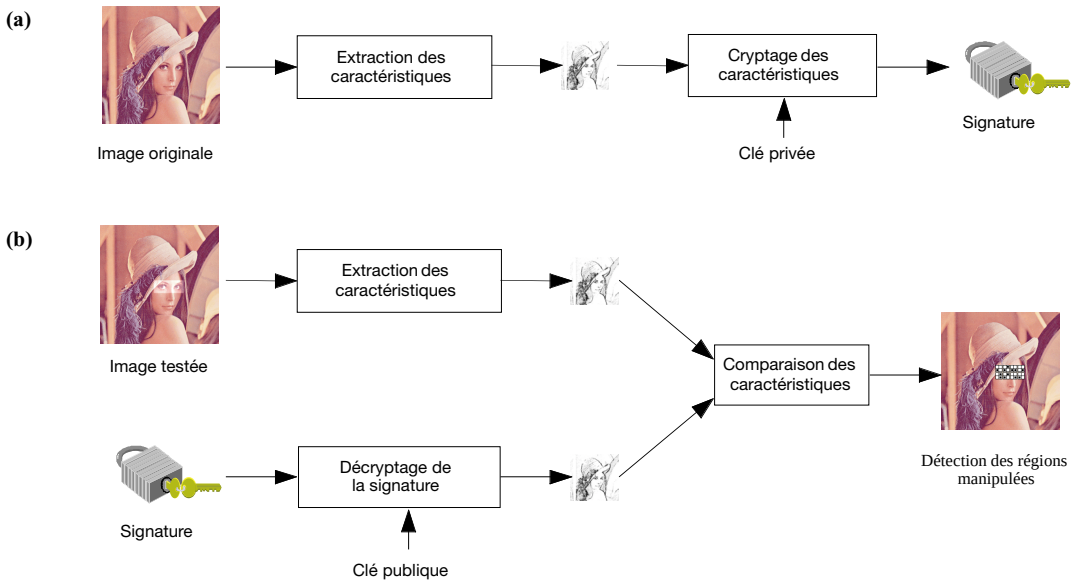

Figure 6 – Principe général d'un système d'authentification utilisant une signature externe (a) Génération de la signature, (b) Vérification de l'intégrité de l'image

2.5. tableau récapitulatif des différentes méthodes

La figure 7, ci-dessous, représente un tableau récapitulatif des différentes méthodes présentées dans cet article. A chaque méthode est associés la catégorie (tatouage fragile, semi-fragile, signature numérique) à laquelle elle appartient, le type de données d'authentification utilisé, le support des données d'authentification, les objectifs en terme d'intégrité (*i.e* stricte ou contenu) et si elle offre une localisation et/ou une reconstruction possible des régions manipulées.

	Catégorie	Marque[1]		Support	Intégrité[2]	Localisation	Reconstruction
Yeung et Mintzer [48]	fragile	logo prédéfini	non	pixels	stricte	oui	non
Walton [42]	fragile	checksums	oui	LSB	stricte	oui	non
Fridrich et Goljan [12]	fragile	image comp.	oui	LSB	stricte	oui	oui
Wong [46]	fragile	hachage	oui	LSB	stricte	oui	non
Lin et Chang [25]	semi-fragile	coef. DCT	oui	DCT	contenu	oui	oui
Wolfgang et Delp [44] (1)	semi-fragile	m-sequences	non	pixels	contenu	oui	non
Rey et Dugelay [36]	semi-fragile	luminance	oui	IFS	contenu	oui	oui
Fridrich [13], [14]	semi-fragile	bloc dep.	oui	pixels	contenu	oui	non
Kundur et Hatzinakos [19]	semi-fragile	bruit aléatoire	non	ondelettes	stricte	oui	non
Lin et Chang [21]	semi-fragile	bruit aléatoire	non	ondelettes	contenu	oui	non
Queluz [32]	signature	contours	oui	externe	contenu	oui	non
Bhattacharjee et Kutter [5]	signature	points intérêts	oui	externe	contenu	oui	non
Lin et Chang [22], [23]	signature	coef. DCT	oui	externe	contenu	oui*	non
Wolfgang et Delp [44] (2)	signature	hachage	oui	externe	stricte	oui*	non

Figure 7 – Tableau récapitulatif des méthodes assurant un service d'intégrité.

1 – indication sur le fait que les données d'authentification sont dépendantes ou indépendantes de l'image.
2 – intégrité : principalement en termes de sensibilité à la compression Jpeg.
*oui : ambiguïté dans la localisation des régions manipulées.

2.6. attaques malveillantes

Avant conclure ce tour d'horizon des méthodes permettant d'assurer un service d'intégrité pour les images, il convient d'aborder le problème des attaques malveillantes de pirates (ou crackers). L'objectif commun de ces attaques, n'est pas de détourner le contenu d'une image (comme les manipulations présentées au paragraphe 1.3), mais d'utiliser les failles ou les faiblesses d'un système d'authentification afin de le tromper, autrement dit faire croire au système qu'une image est intègre alors que son contenu a été modifié (ou l'inverse dans certains cas). Ce paragraphe n'a pas pour but de faire l'inventaire de toutes ces attaques, mais d'en présenter quelques unes parmi les plus fréquentes. Même si certaines de ces attaques paraissent triviales et simples à prévenir, il est néanmoins très important d'en tenir compte lors de l'élaboration d'un algorithme d'authentification.

Une des attaques les plus courantes contre les systèmes à base de tatouage fragile, consiste à tenter de modifier une image protégée sans affecter le tatouage qu'elle contient, ou bien encore à tenter de créer une nouvelle marque que le détecteur considérera comme authentique. Prenons par exemple le cas volontairement simplifié où l'intégrité d'une image est assurée par une marque fragile, indépendante du contenu, insérée dans les LSB des pixels. Il est clair que si on modifie l'image sans se préoccuper de savoir quels sont les bits affectés par la manipulation, on a toutes les chances pour que la marque soit dégradée et l'attaque détectée. Par contre, si on prend soin de modifier l'image sans toucher aux LSB, la marque restera intacte et le système ne décèlera aucune falsification.

D'un point de vue plus général, dès lors que l'intégrité est assurée par une marque indépendante du contenu de l'image à protéger il est possible d'imaginer une attaque qui recopie une marque valide d'une image dans une autre (exemple : la « Copy Attack » de Kutter *et al.* [20]). De cette manière la deuxième image se retrouve alors protégée. Cette attaque peut très bien être appliquée sur la même image ; dans ce cas, la marque est dans un premier temps retirée de l'image, l'image est ensuite manipulée, et enfin la marque est réinsérée dans l'image manipulée, trompant ainsi le système d'authentification.

Dans le même esprit, la « Collage-Attack » [15] [28] proposée par Fridrich *et al.* consiste à créer une image contrefaite de toutes pièces à partir d'une banque d'images protégées par la même marque et la même clé. Cette attaque ne présuppose aucune connaissance *a priori* sur la marque binaire cachée, ni sur la clé secrète utilisée. Son principe est relativement simple puisqu'il consiste à remplacer chaque pixel de l'image à manipuler par le pixel qui lui est le plus similaire parmi les pixels de même position des images de la base. La difficulté de cette méthode est de disposer d'une banque d'images suffisamment variées pour obtenir une image falsifiée de bonne qualité visuelle.

Une autre attaque classique consiste à essayer de trouver la clé secrète utilisée pour générer la marque. Ce type d'attaque, également appelé « Brute Force Attack », est très connu par la communauté « sécurité ». Une fois la clé trouvée, il devient alors très facile pour un pirate de falsifier la marque d'une image protégée avec cette clé. La seule parade efficace est d'utiliser des clés de grande taille de manière à rendre cette attaque très dissuasive en termes de temps de calcul.

3. remarques concluantes

Le volume sans cesse croissant d'échange de données numériques engendre de nouveaux besoins en termes de sécurité des informations. Les documents multimédia, et les images en particulier, n'échappent pas à ce phénomène. Les utilisateurs attendent des solutions performantes permettant notamment d'assurer la protection des droits d'auteur, mais également de garantir l'authenticité des documents multimédia. Cette demande est d'autant plus forte que les techniques de manipulation d'image sont de plus en plus sophistiquées et accessibles au plus grand nombre, et que les exemples de falsifications de documents deviennent malheureusement de plus en plus fréquents. Dans ce contexte, le tatouage d'image, bien qu'étant un domaine de recherche relativement récent, peut apporter des éléments de réponse complémentaire aux méthodes de cryptographie classiques, en proposant notamment une approche privilégiant une intégrité en terme de contenu, à une intégrité numérique stricte.

Cependant, dans l'état actuel des recherches, il est difficile d'affirmer quelle approche semble la plus appropriée à assurer un service d'intégrité adapté aux images et d'une manière plus générale aux documents multimédia. Il n'existe pas, pour l'instant, de solution répondant parfaitement à ce problème. Les méthodes reposant sur un tatouage fragile, sont très sensibles à la moindre altération de l'image, n'offrant par conséquent qu'un service d'intégrité stricte, relativement éloignée des besoins des utilisateurs. Néanmoins, les techniques de tatouages fragiles ont l'avantage, par rapport aux méthodes classiquement utilisées en sécurité, de permettre une localisation précise des régions qui ont été manipulées (dans le cas où, bien évidemment, l'image n'a pas subi en plus de manipulation globale). La tendance actuelle s'oriente, cependant de plus en plus vers l'utilisation de méthodes dites semi-fragiles. Ces méthodes sont beaucoup plus tolérantes vis-à-vis des manipulations bienveillantes, telles qu'une compression Jpeg de bonne qualité. Cette souplesse est rendue possible en partie grâce à des algorithmes de tatouage à robustesse ciblée (*i.e.* la marque n'est résistante qu'à certaines manipulations bien déterminées), mais aussi par l'utilisation de données d'authentification de haut niveau, basées sur le contenu sémantique de l'image plutôt que sur les valeurs numériques des pixels. L'utilisation d'une marque dépendante du contenu de l'image permet, d'une part d'accroître la robustesse de la méthode vis-à-vis d'attaques malveillantes comme la « Collage Attack », et d'autre part, en fonction des caractéristiques choisies, une éventuelle réparation partielle des régions altérées.

Par ailleurs, les méthodes ayant recours à une signature numérique externe offrent une alternative intéressante aux techniques classiques de tatouage d'image, dans la mesure où il n'y a plus vraiment de limitation en terme de capacité, ni de problème de robustesse, offrant ainsi une possibilité de localisation plus fine des régions manipulées, une restauration de meilleure qualité, et un risque de fausses alarmes réduit. Ce type de méthodes disposent déjà d'une bonne expertise de la part de la communauté sécurité, cependant, elles ne sont pas sans défaut. De nouvelles contraintes sont à prendre en considération, comme par exemple garantir l'intégrité de la signature, ainsi que celle du couple image/signature). De plus elles s'avèrent également peu pratiques d'utilisation pour des documents multimédia. Enfin, il n'est pas exclu d'imaginer à l'avenir des méthodes basées sur une coopération tatouage robuste et signature externe, où le tatouage ne serait qu'un identifiant permettant d'accéder à la signature enregistrée auprès d'une tierce personne de confiance [3].

Avant de conclure, il est intéressant de remarquer, que malgré les imperfections des méthodes existantes permettant d'assurer un service d'intégrité d'image, des produits commerciaux, logiciels et matériels, sont d'ores et déjà disponibles pour le grand public. Les principaux sont : le système DSS de Kodak [18] (Digital Signature Standard, norme reconnue par le *National Institute of Standards and Technology* [42]), le système IAS (Image Authentication System) d'Epson [10], Veridata de Signum Technologies [41], Eikonamark d'Alpha-Tec Ltd [2], Mediasign de MediaSec [27] et PhotoCheck d'AlpVision [1]. Les systèmes proposés par Kodak et Epson sont directement intégrés au niveau de leurs appareils photo numériques afin de protéger les images dès leurs acquisitions. Les applications visées par ces différents produits sont multiples. Cela va de l'intégrité des images à des fins d'expertises, à la protection de documents numériques (images de caméras de vidéo surveillance par exemple) pouvant être utilisés comme éléments de preuve lors de procès. AlpVision et Signum Technologies proposent des utilisations plus originales comme par exemple renforcer la sécurité de documents papiers tels qu'un passeport ou un badge d'accès en tatouant la photo d'identité.

BIBLIOGRAPHIE

[1] AlpVision. http://www.alpvision.com

[2] Alpha-Tec Ltd. http://www.alphatecltd.com

[3] RNRT – AQUAMARS. http://www.telecom.gouv.fr/rnrt/wprojets.htm

[4] IST – CERTIMARK: a benchmark suite for watermarking of visual content and a certification process for watermarking algorithms. http://www.certimark.org

[5] S. Bhattacharjee and M. Kutter. Compression Tolerant Image Authentication. *IEEE International Conf. on Image Processing (ICIP'98)*, Chicago, USA, Oct. 1998.

[6] Ça m'intéresse. Pourquoi faut-il se méfier des images ?, No 235, Sept. 2000.

[7] J.-L. Dugelay & S. Roche. Introduction au tatouage d'images. *Annales des Télécommunications*, 54, no 9-10, pp. 427-437, 1999.

[8] J.-L. Dugelay. Procédé de dissimulation d'informations dans une image numérique. *Brevet INPI FR 98-04083 (EURECOM 09-FR)*, March 1998.

[9] J.-L. Dugelay & S. Roche. Process for marking a multimedia document, such an image, by generating a mark. *Pending patent EP 99480075.3* (EURECOM 11/12 EP), July 1999.

[10] Epson. http://www.epson.co.uk/

[11] J. Fridrich. Robust Bit Extraction From Images. *ICMCS'99*, Florence, Italy, june 1999.

[12] J. Fridrich and M. Goljan. Protection of Digital Images using Self Embedding. *The Symposium on Content Security and Data Hiding in Digital Media*, New Jersey Institute of Technology, Mar. 1999.

[13] J. Fridrich. Image Watermarking for Tamper Detection. *Proceedings IEEE Int. Conf. on Image Processing (ICIP'98)*, Chicago, USA, Oct. 1998.

[14] J. Fridrich. Methods for detecting changes in digital images. Proceedings IEEE *Int. Conf. on Image Processing (ICIP'98)*, Chicago, USA, Oct. 1998.

[15] J. Fridrich, M. Goljan & N. Memon. Further Attacks on Yeung-Minzer Fragile Watermarking Scheme. . *SPIE International Conf. on Security and Watermarking of Multimedia Contents II*, vol. 3971, No 13, San Jose, USA, Jan. 2000.

[16] S. Katzenbeisser & Fabien A. P. Petitcolas, with contributions of: Fabrizio Marongui Buonaiuti, Scott Craver, Jean-Luc Dugelay, Frank Hartung, Neil F. Johnson, Martin Kutter, Jong-Hyeon Lee, Stanley Lai, Adrian Perrig, Stéphane Roche. Information hiding techniques for steganography and digital watermarking. Artech House Books, ISBN 1-58053-035-4, Dec. 1999. http://www.cl.cam.ac.uk/~fapp2/papers/book99-ih/

[17] A. Kerckhoffs. La cryptographie militaire. Journal des Sciences militaires, 9, pp. 5-38, 1883.

[18] Kodak. Understanding and Intergrating KODAK Picture Authentication Cameras. (White paper) http://www.kodak.com/US/en/digital/software/imageAuthentication/

[19] D. Kundur and D. Hatzinakos. Towards a Telltale Watermarking Technique for Tamper-Proofing. *IEEE International Conf. on Image Processing (ICIP'98)*, Chicago, USA, Oct. 1998.

[20] M. Kutter, S. Voloshynocskiy and A. Herrigel. The Watermark Copy Attack. *In Proceedings of SPIE Security and Watermarking of Multimedia Content II*, vol. 3971, San Jose, USA, Jan. 2000.

[21] C.-Y. Lin and S.-F. Chang. A Watermark-Based Robust Image Authentication Using Wavelets. *ADVENT Project Report*, Columbia University, Apr. 1998.

[22] C.-Y. Lin and S.-F. Chang. Generating Robust Digital Signature for Image/Video Authentication. *Multimedia and Security Workshop at ACM Multimedia 98*, Bristol, UK, Sep 1998.

[23] C.-Y. Lin and S.-F. Chang. A Robust Image Authentication Method Surviving JPEG Lossy Compression. *SPIE Storage and Retrieval of Image/Video Database*, San Jose, USA, Jan. 1998.

[24] C.-Y. Lin and S.-F. Chang. Distorsion Modeling and Invariant Extraction for Digital Image Print-and-Scan Process. *ISMIP 99*, Taipei, Taiwan, Dec. 1999.

[25] C.-Y. Lin and S.-F. Chang. Semi-Fragile Watermarking for Authenticating JPEG Visual Content. *SPIE International Conf. on Security and Watermarking of Multimedia Contents II*, vol. 3971, No 13, San Jose, USA, Jan 2000.

[26] B.S. Manjunath, C. Shekhar and R. Chellappa. A new approch to image feature detection with applications. *Pattern Recognition*, 29(4) : 627-640, 1996.

[27] Mediasec. http://www.mediasec.com

[28] N.D. Memon and J. Fridrich. Further Attacks on the Yeung-Mintzer Fragile Watermark, *SPIE International Conf. on Security and Watermarking of Multimedia Contents II*, vol. 3971, No 13, San Jose, USA, Jan 2000.

[29] C. Nastar. Indexation d'Images par le Contenu : un Etat de l'Art. *Coresa 97*, Issy-les-Moulinaux, France, Mar. 1997.

[30] W.B. Pennebaker and J.L. Mitchell. JPEG still image data compression standard. *Van Nostrand Reinhold Compagny*, 1992.

[31] J.G. Proakis, Digital Communications, *Third Edition*, *McGraw Hill*, New York, 1995.
[32] M.P. Queluz. Towards Robust, Content Based Techniques for Image Authentication. IEEE *Signal Processing Society 1998 Workshop on Multimedia Signal Processing*, Dec. 1998.
[33] K.R. Rao and P. Yip. Discrete cosinus transform : algorithms, advantages, applications. *Academic Press Inc*., 1990.
[34] R. L. Rivest, A. Shamir & L. Adelman. On Digital Signatures and Public Key Cryptosystems. *MIT Laboratory for Computer Science Technical Memorandum 82*, Apr. 1977.
[35] Agence Reuters http://www.reuters.com
[36] C. Rey & J.-L. Dugelay. Blind Detection of Malicious Alterations On Still Images Using Robust Watermarks. *IEE Secure Images and Image Authentication colloquium*, London, UK, Apr. 2000.
[37] E. Roskis. Images truquées. *Le Monde Diplomatique*, Jan. 1995.
[38] J.J.K. Ó Ruanaidh and T. Pun. Rotaion, Scale and Translation Invariant Digital Image Watermarking. *Proc. of the ICIP*, vol. 1, pp. 536-539, Santa Barbara, California, Oct. 1997.
[39] R.G. Van Schyndel, A. Z. Tirkel, and C. F. Osborne. A Digital Watermark. *Proceedings IEEE Int. Conf. on Image Processing (ICIP'94)*, Vol. 2, pp. 86-90, Austin, Texas, Nov. 1994.
[40] SHA-1, Secure Hash Standard (SHS), spécification (FIPS 180-1), http://www.itl.nist.gov/fipspubs/fip180-1.htm
[41] Signum Tecnologies. http://www.signumtech.com/
[42] U.S. Department of Commerce, National Institute of Standards and Technology (FIPS PUB 186-1), http://www.itl.nist.gov/fipspubs/by-num.htm
[43] S. Walton. Information Authentification for a Slippery New Age. *Dr. Dobbs Journal*, vol. 20, No. 4, pp. 18-26, Apr. 1995.
[44] R.B. Wolfgang and E. J. Delp. A watermark for digital images. *Proceedings of the 1996 International Conference on Image Processing*. Vol. 3, pp. 219-222, Lausanne, Switzerland, Sept. 1996.
[45] R.B. Wolfgang and E. J. Delp. Fragile Watermarking Using the VW2D Watermark. *SPIE International Conf. on Security and Watermarking of Multimedia Contents*, vol. 3657, No. 22, EI '99, San Jose, USA, Jan. 1999.
[46] P. Wong. A watermark for image integrity and ownership verification. *Final Program and Proceedings of the IS&T PICS 99*, pp. 374-379, Savana, Georgia, Apr. 1999.
[47] M. Wu and B. Liu. Watermarking for Image Authentication. *IEEE International Conf. on Image Processing*, Chicago, USA, Oct 1998.
[48] M.M. Yeung and F. Mintzer. An Invisible Watermarking Technique for Image Verification. *IEEE International Conf. on Image Processing*, Santa Barbara, USA, Oct. 1997.

Manuscrit reçu le 18 avril 2001

LES AUTEURS

Jean-Luc DUGELAY

Docteur en informatique, Jean-Luc Dugelay est actuellement professeur à l'Institut Eurécom (Sophia Antipolis) au sein du département Communications Multimédia et chercheur invité à l'Université de Californie, Santa Barbara. Ses domaines de recherche sont le traitement, codage, tatouage et indexation des images, les communications vidéo, la réalité virtuelle et l'imagerie 3D. Membre du comité scientifique de l'IEEE Multimédia, il est éditeur associé de plusieurs revues nationales et internationales.

Christian REY

Christian Rey est doctorant au sein du département Communications Multimédia de l'Institut Eurécom. Son sujet de thèse porte sur les mécanismes de sécurité liés à la diffusion des images, principalement sur des applications pour la protection des droits d'auteur et de l'intégrité des images.

Jean-Luc Dugelay et Christian Rey sont co-auteurs de plusieurs brevets, publications et démonstrations dans le domaine du tatouage d'image. Eurécom participe au projet européen IST Certimark [4].