

Schéma conjoint de tatouage et compression des LDI(s) pour des flux auto-stéréoscopiques

N. Khelifi¹, Z. Ahmed Foitih² et L. Lucas³

¹LRIIR, Faculté des Sciences Exactes et Appliquées Université d'Oran - Algérie

²Département d'Electronique - Faculté du Génie Electrique USTO Mohamed Boudiaf - Oran - Algérie

³CReSTIC SIC EA3804 - Université de Reims Champagne-Ardenne

Résumé

Dans cet article est présenté un schéma conjoint de tatouage et compression des LDI(s) (Layer Depth Image) générées à partir d'un jeu d'images multi-vues. Chaque calque ou "layer" correspond à une vue et contient les informations qui ne sont pas contenues dans les autres calques. Les calques de profondeur sur lesquels nos travaux se basent contiennent uniquement des informations non redondantes (ou résiduelles). Elles sont extraites à partir de n vues d'entrée et de leur carte de disparité et correspondent à des zones occultées. Les pixels occlus, qui sont susceptibles d'être visualisés à partir d'autres angles de vue, se voient conférer une plus forte protection contre les erreurs de transmission ou de compression. Nous cherchons donc à effectuer l'insertion d'une marque dans un ensemble de LDI(s) pendant la phase de quantification et de compression à l'aide d'une méthode qui doit tenir compte de ces pixels occlus (ou résiduels) localisés dans des zones dispersées de chaque layer. Notre approche de tatouage joint à la compression est basée sur la quantification vectorielle algébrique à zone morte (QVAZM). Cette technique permet d'effectuer conjointement compression et tatouage à l'aide de la QVAZM en réalisant simultanément la quantification et l'insertion de la marque durant la compression.

Mots clé : Images auto-stéréoscopiques, Cartes de profondeur, Layer Depth Image, Sécurité, Tatouage numérique, Quantification vectorielle à zone morte, Compression

1. Introduction

Le nombre croissant d'applications utilisant les technologies multimédias numériques a fortement accentué la nécessité d'assurer la protection de ces médias notamment en termes de droit d'auteur. Nous nous intéressons ici à l'une de ces techniques, le tatouage numérique d'images ou de vidéos, qui permet d'insérer de manière imperceptible par l'utilisateur final un message (copyright, message de vérification, ...) à un fichier hôte. Ce message, généralement appelé marque, se présente comme un ensemble de bits dont le contenu dépend de l'application. De nombreuses méthodes d'insertion existent dans la littérature [CMB*08a]. De manière générale, la détection de la marque doit être fiable, sans fausse détection et si possible pas de faux rejet. En outre, les marques doivent être :

- sécurisées - une modification non autorisée comme le retrait de la marque doit être impossible ;
- insensibles - il doit être possible de détecter et de décoder la marque sans l'utilisation de l'image ou de la vidéo d'origine ;
- perceptuellement invisibles et ne doivent pas dégrader la qualité d'image ou de vidéo ;

- résistantes aux manipulations telles que les transformations photométriques (e.g. correction de filtrage ou luminance), transformations géométriques (e.g. translation, rotation, redimensionnement, recadrage ou changement aspect ratio), conversions analogiques-numériques (resp. numériques-analogiques), la numérisation, la compression avec perte et autres attaques cryptographiques.

Dans le cadre de la vidéo et de surcroît de la vidéo 3D [KCA10, MCP13, BBD14], la marque doit résister à différents schémas de compression vidéo, ne pas altérer l'information de profondeur et être détectable n'importe où dans le film y compris sur un court laps de temps.

Ce article présente une idée de schéma conjoint de tatouage et de compression sur des images type LDIs, générées à partir d'un jeu d'images auto stéréoscopiques, prises à partir de n points de vue ($n = 8$). Notre travail s'intègre dans un processus élaboré par [Bat12, Niq11]. Aussi, dans la première partie de cet article, seront définies les LDI(s) et les différentes étapes de traitement qui leur ont été appliquées et qui rentrent dans le processus de compression. C'est sur ces LDIs, particulièrement les calques résiduels que va être appliqué un schéma d'insertion d'une marque. Cette étape d'insertion s'applique au moment de la compression des calques. Après avoir subi une transformation en ondelettes, ces images seront quantifiées par la méthode de Quan-

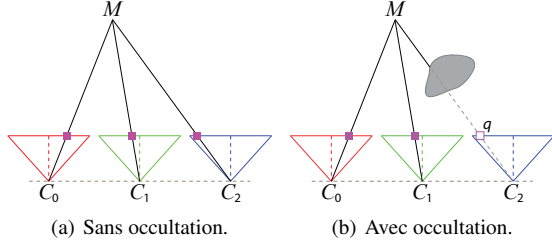


Figure 1: Soit M un point 3D et son match correspondant m (■ pixel $\in m$, □ pixel $\notin m$). Sur la figure (b), le point M est caché par l'objet gris et n'est pas visible depuis la vue 2. Par conséquent, le pixel $q = (2, m_x - 2m_d, m_y)$ dans la vue 2 ne correspond pas à la projection de M et ne fait donc pas partie de m même s'il satisfait à l'équation 2.1. Ainsi, nous dirons que M est visible dans les vues 0 et 1 et non visible dans la vue 2.

tification Vectorielle Algébrique à Zone Morte qui permet de localiser les zones de forte activité représentées par les vecteurs d'ondelettes dits significatifs et qui sont de bons candidats au tatouage. Cela permet d'améliorer la robustesse du tatouage. L'insertion de la marque se fait donc sur ces vecteurs significatifs, les vecteurs seuils (ou nuls) sont exclus du processus de tatouage.

2. Génération de LDI

2.1. Estimation des disparités

La plupart des problèmes rencontrés par les méthodes de compression de vidéo 3D basées LDV (Layer-Depth-Video) sont dus à des incohérences entre les cartes de disparité le plus souvent présentes dans les zones d'occlusion. Pour remédier à cela, nous nous basons dans cet article, sur des algorithmes d'estimation de disparité fonctionnant en arithmétique entière de sorte à ce que tous pixels correspondants aient la même ordonnée dans leur base d'image respective. L'approche retenue est celle de [NPR10a] qui permet à la fois de produire des disparités cohérentes et de lever toutes ambiguïtés.

Soit P l'ensemble des pixels contenus dans les n vues. On définit $p = (i, p_x, p_y)$ comme le pixel de coordonnées (p_x, p_y) de la vue i et m comme un match (un ensemble de pixels dont les projections ont pour origine le même point 3D M - voir Figure 1(a)). Soit (x_0, y_0) , les coordonnées du pixel $\in m$ dans la première vue (la vue 0). Dans le cadre d'une géométrie de capture simplifiée (distribution linéaire équidistantes des centres optiques), les coordonnées des autres pixels de m dans la base induite par les autres vues sont alors faciles à trouver. Par exemple, la position du pixel dans la vue k est donnée par

$$\begin{pmatrix} x_k \\ y_k \end{pmatrix} = \begin{pmatrix} x_0 - k \times m_d \\ y_0 \end{pmatrix},$$

où m_d est une valeur de disparité entière associée à m .

L'utilisation de disparité entière assure que l'expression $x_0 - k \times m_d$ correspond à une coordonnée entière, ce qui évite tous problèmes d'ambiguïté dans le processus d'extraction des différents calques de disparité. Par convention, nous

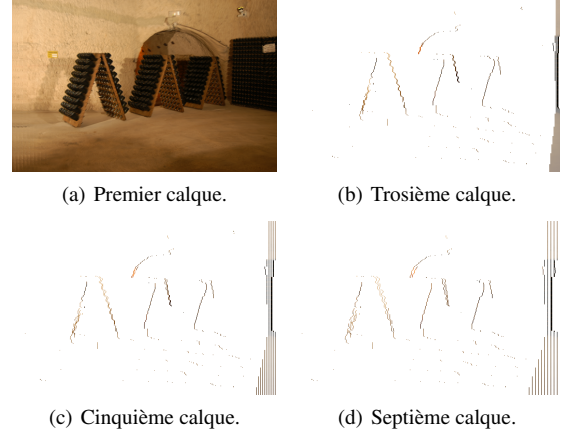


Figure 2: Calques générés à partir de la séquence "Ruinar".

posons $m_x = x_0$ et $m_y = y_0$, définissant la position de la projection de M dans la première vue (la vue 0). Ainsi, un pixel p de la vue k peut appartenir à un match m si et seulement si la condition suivante est vraie :

$$\begin{cases} p_x = m_x - k \times m_d \\ p_y = m_y \end{cases}$$

L'algorithme d'estimation de disparité vise à déterminer une partition de tous les pixels dans les matchs. Cette partition est un ensemble de correspondances non vides (contenant chacun au moins un pixel) qui n'ont aucun pixel en commun. De plus, l'union de tous les matchs de la partition est égale à l'ensemble P de tous les pixels. Une partition g vérifie donc les conditions suivantes :

$$\begin{aligned} \forall m \in g, m &\neq \emptyset \\ \forall m, n \in g, m \cap n &= \emptyset \\ \bigcup_{m \in g} m &= P \end{aligned}$$

Dans la pratique, cet algorithme commence par initialiser la partition avec les matchs correspondant aux plus grandes valeurs de disparité. Toutes ces valeurs de disparité sont ensuite analysées des plus élevées aux plus petites. Pour chaque valeur de disparité α , tous les matchs déjà présents dans la partition sont pris en compte et nous cherchons à évaluer s'ils existent vraiment ou pas. Si un match n'existe pas, il est alors retiré de la partition et remplacé par de nouveaux matchs de disparité α' plus faibles (a priori donc visibles). L'ensemble de ces contraintes sont globalement maintenues cohérentes au moyen d'un algorithme de graph-cuts [BVZ99, KZ04]. Pour plus de détails sur cette approche, il est possible de se référer à [NPR10b].

2.2. Extraction des calques de disparités

Le principe de mise en correspondance présenté précédemment permet de détecter l'ensemble des redondances. Tous les pixels appartenant à un même match sont des projections d'un point 3D unique de la scène et sont considérés comme redondants. Supprimer ces redondances revient donc à ne maintenir qu'un seul pixel de chaque match de la partition (celui ayant le plus petit ordre dans l'image multi-vues).

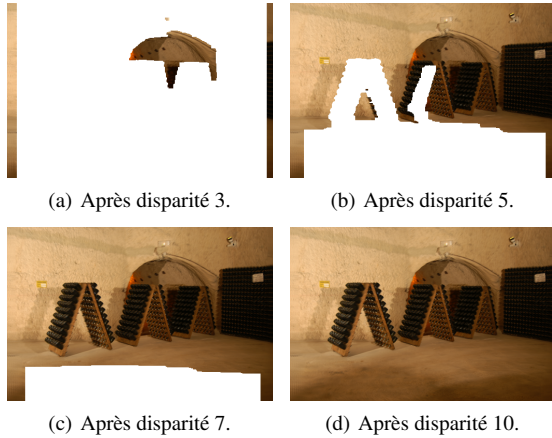


Figure 3: Reconstruction de l'image à partir de calques.

Ainsi, le premier calque conservera tous les pixels, le second ne conservera que les pixels ne pouvant être déduits du premier calque, le troisième conservera uniquement les pixels ne pouvant être déduits ni du premier ni du second calque et ainsi de suite. Notre définition du LDI diffère de celle habituellement décrite dans la littérature [YLKYS07, JMG09] : dans notre cas, le nombre de calques est toujours égal à n et chacun est attaché à sa vue correspondante (et pas à une vue de référence unique). En outre, le rapport de pixels est approximativement constant d'un calque à l'autre, à l'exception du premier, bien entendu. Afin de réduire les artefacts dus aux variations de lumière d'une image à l'autre, seule la couleur moyenne de tous les pixels d'un match est conservée. La figure 2 illustre ce procédé d'extraction des différents calques sur la séquence "Ruinart".

2.3. Reconstruction de l'image

Le processus de reconstruction des différentes vues consiste s'appuyer pour partie sur l'algorithme du peintre [NNS72]. Il exploite également la contrainte exposée précédemment, contrainte de cohérence qui assure qu'un match ne peut être visible dans une autre image que si celui-ci est obstrué par un autre match avec un écart supérieur. L'idée consiste donc à "étaler" les pixels par ordre de disparité croissante dans les n . Certains pixels pourront être dessinés sur d'autres déjà existants de sorte à reproduire les zones occlusions présentes dans la scène. La figure 3 illustre plusieurs étapes de cette phase de reconstruction d'une image à partir des calques de la Figure 2.

3. Compression des LDI(s)

Nous avons présenté dans la précédente section une structure de représentation d'images multi-vues appelée LDI. Cette structure permet au moyen de calques de stocker pour chaque pixel non seulement sa couleur mais également sa profondeur relativement au calque qui le porte. Elle est générée à partir d'un jeu d'images multi-vues. Les n calques (ou "layers") sont associées à chacune des n vues. La première couche, appelée couche de référence, contient tous les pixels de la première vue, alors que les $n - 1$ couches suivantes

(désignées par le terme "couches résiduelles") contiennent uniquement l'information n'appartenant pas à la première couche (à cause des zones d'occultations).

La compression d'images multi-vues [Bat12] basée sur ce principe repose sur les étapes suivantes :

- Etape 1 - Conversion colorimétrique de l'espace RGB vers l'espace YCbCr.
- Etape 2 - Application de la DWT sur chacune des $n - 1$ calques : ces calques ne contenant que des pixels résiduels localisés dans des zones dispersées, la SA-DWT (Shape-Adaptative Discrete Wavelet Transform) [LL00] a été utilisée pour ne compresser que certaines zones d'une image tout en omettant le reste de l'information. La sélection de ces zones a été faite par masque binaire lui indiquant si le pixel est à prendre en considération pour la compression ou pas.
- Etape 3 - Codage des différents coefficients d'ondelettes par application de l'algorithme SPIHT associé à un codeur arithmétique binaire.

Notre objectif est de présenter un schéma de tatouage conjoint à la compression pour insérer une marque dans ces calques résiduelles en choisissant une méthode qui tienne compte des particularités de ces LDI. La méthode qui nous a semblé la plus appropriée à ce type d'images est celle qui utilise la quantification vectorielle algébrique à zone morte (QVAZM) et d'insérer la signature sur des coefficients quantifiés de façon à répartir les bits de la signature sur les différents calques.

4. Le tatouage joint à la compression

La compression jointe au tatouage a suscité un intérêt récent. Associer tatouage et compression permet d'améliorer la capacité d'insertion du tatouage et d'assurer une meilleure détection tout en maintenant une bonne qualité d'image. La compression représente, non seulement un passage obligé du stockage ou du transfert d'images, mais aussi l'une des attaques les plus destructrices vis-à-vis du tatouage.

Le domaine multi-résolution est la base des normes de compression avec perte. Son utilisation comme support d'insertion de la signature assure une meilleure robustesse vis-à-vis de cette norme de compression. L'ondelette 9/7 "Daubechies" utilisée, permet le passage du domaine spatial vers le domaine multi-résolution.

Lorsque la marque est insérée directement durant le processus de codage, celle-ci devient partie intégrante de l'image. Pour ce faire, nous intégrons l'étape d'insertion dans le schéma de codage basé sur la quantification vectorielle algébrique avec zone morte afin d'en exploiter les principaux avantages, à commencer par son aptitude à localiser les zones de forte activité. Ces zones sont représentées par les vecteurs d'ondelettes dits "significatifs" qui sont donc de bons candidats au tatouage. Aussi, la phase d'insertion du message se situe pendant l'étape 2 de l'algorithme de compression sus cité (cf. section 3) à partir de l'application de la DWT sur les $n - 1$ calques résiduels. L'application du masque binaire et de la SA-DWT sur les calques laisse place à la quantification algébrique à zone morte. Les quelques vecteurs quantifiés sélectionnés seront alors tatoués.

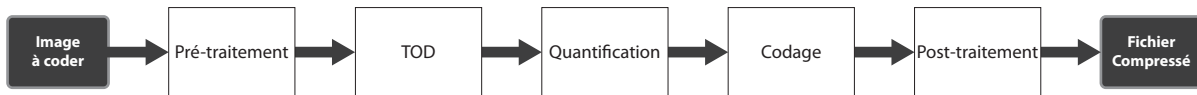


Figure 4: Schéma général de la chaîne de compression d'une image.

Le tatouage d'image est divisé en deux parties : la première est la phase d'insertion de la marque, la deuxième est la phase de détection. Nous ne considérerons, dans cet article, que la phase d'insertion.

4.1. Méthode de tatouage proposée

Le schéma de tatouage que nous proposons se base sur un schéma substitutif dans le domaine multi-résolution où les coefficients d'ondelettes ont subi une Quantification Vectorielle Algébrique à Zone Morte (QVAZM) [Voi03]. L'image obtenue est donc une image comprimée et tatouée. En d'autres termes, nous appliquons notre technique de tatouage directement sur les données quantifiées. Différents travaux ont montré l'intérêt d'utiliser ce type de schéma [KH98, CMB*08b]. L'objectif commun de la compression et du tatouage d'images est de minimiser l'impact visuel des modifications apportées à l'image originale. Cependant, les modifications apportées ont des objectifs totalement opposés : la compression d'images vise à supprimer les données imperceptibles afin de réduire la quantité d'information, alors que le tatouage consiste à ajouter une information imperceptible (intégration d'une signature).

Ce schéma de tatouage (cf. Figure 5) sera appliqué sur l'ensemble des calques résiduels contenant des pixels correspondants aux zones d'occlusions. La particularité de ce schéma est que la marque (ou le message) sera étendu sur les différentes couches résiduelles. De ce fait, la première étape consiste à sélectionner les vecteurs de coefficients d'une seule sous bande par couche qui vont porter une partie de la signature et la deuxième permet l'insertion de cette dernière sur les coefficients de ces sous bandes de marquage suivant une fonction d'insertion donnée.

La justification de ce schéma de tatouage se base sur aspects suivants :

- Le domaine multi-résolution : c'est le domaine le mieux adapté au système visuel humain, dans le sens où celui-ci est plus ou moins sensible à certaines gammes fréquentielles sélectionnées. La décomposition en sous-bandes de la transformée en ondelettes permet d'isoler les fréquences adaptées au tatouage. De plus, c'est le domaine utilisé par le standard de compression SPIHT déjà utilisé dans notre cas. Le niveau de décomposition représentant le niveau d'insertion est choisi de telle sorte que les coefficients de ce niveau soient généralement bien préservés lors de la phase de quantification. Généralement, c'est ceux qui contiennent les informations les plus significatives de l'image, cela permet de mieux préserver la signature

face à différentes attaques. De plus la modification de certains coefficients de ce niveau n'engendre pas de dégradations trop importantes lors de la phase de reconstruction de l'image. Ce niveau se situe entre les basses et moyennes fréquences.

- Un schéma substitutif : ce type de schéma, contrairement au schéma additif, permet d'utiliser au maximum la capacité du support d'insertion. Par exemple, il est possible d'insérer 1 bit d'information issu du message par vecteur (non nul) d'une sous-bande d'un calque.
- La QVAZM : la Quantification Vectorielle Algébrique avec Zone Morte d'une sous-bande présente l'avantage de ne conserver que les coefficients d'ondelettes significatifs, les autres étant mis à zéro. Elle permet ainsi d'améliorer la robustesse de la signature insérée face à différents types d'attaques [LL00]. Le choix de la QVAZM, dans ce cas, est justifié d'une part par le fait que les couches de la LDI contiennent peu d'informations correspondants aux pixels occlus et que ces informations sont dispersées et d'autre part parce que les différentes sous bandes sont considérées indépendantes les unes des autres et de ce fait, la modification de certaines d'entre elles, par insertion d'une marque, ne perturbe pas les autres (approche intra bande). En plus, le choix des sous bandes peut être différent d'un calque à l'autre.

4.2. L'opération d'insertion

L'opération d'insertion se fait de la même manière sur les différents calques d'une LDI à l'exception d'une sélection différente des vecteurs à tatouer. Elle consiste en fait à sélectionner des vecteurs des sous bandes puis à insérer les bits de la signature sur les différents vecteurs des différentes couches de la LDI.

Les étapes d'insertion sont présentées pour un calque de la LDI et doivent être répétées sur les autres calques :

- La phase de Quantification. Le schéma de quantification est réalisé comme suit : une sous bande SB_k est sélectionnée. De cette sous bande sont sélectionnés les vecteurs susceptibles d'être tatoués et vont donc être quantifiés par la QVAZM : (vecteurs quantifiés Y_i).
- La phase de tatouage : A l'aide de la clé K , vont être sélectionnés les vecteurs quantifiés qui vont porter la signature S . Le tatouage est réalisé par un schéma de substitution. Il permet d'insérer 1 ou quelques bits de la signature $w = (b_0, \dots, b_{M_w})$ sur des vecteurs quantifiés Y_i de la sous-bande SB_k . A l'aide de la clé K , vont être sélectionnés les vecteurs $S(Y_i, K)$ qui vont porter la signature. L'ensemble de ces vecteurs est noté : $S(Y_i, K) = Y_s$.

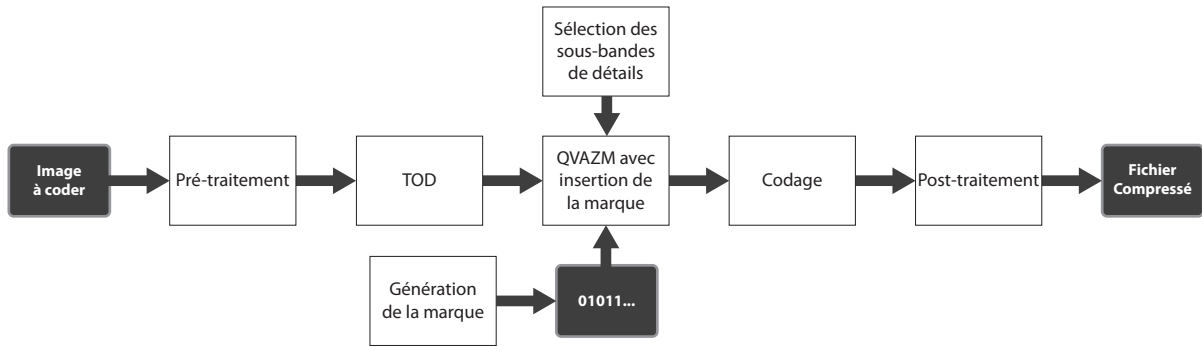


Figure 5: Schéma général du tatouage d'une image joint à la compression.

Les vecteurs qui n'ont pas été sélectionnés sont notés $Y_{\bar{S}}$. Nous avons donc les propriétés suivantes.

$$Y_i = Y_S \cup Y_{\bar{S}}$$

$$Y_S \cap Y_{\bar{S}} = 0$$

Si l'on considère qu'un bit b_j de la signature est inséré par vecteur sélectionné Y_S alors ce vecteur tatoué s'écrira : $W_b(Y_{s_j}, b_j) = Y_{s_j}^w$ (vecteur sélectionné tatoué avec le bit b_j de la marque).

L'étape d'insertion modifie les composantes du vecteur Y_{S_j} en fonction du bit b_j à insérer. Appliquée à tous les vecteurs sélectionnés, l'étape d'insertion conduit à l'ensemble des vecteurs tatoués Y_S^w .

$$W_s(Y_s, b) = Y_s^w \text{ avec } Y_s^w = \{Y_{s_j}^w\}$$

Les vecteurs sélectionnés et tatoués Y_S^w vont être ensuite intégrés aux vecteurs non sélectionnés $Y_{\bar{S}}$ pour former l'ensemble des vecteurs Y_w de la sous-bande tatouée SB_k^w . Ces différentes étapes réalisent la substitution $Sub(Y_s, Y_s^w, K)$ des vecteurs avec les vecteurs tatoués Y_s^w afin d'intégrer la signature. Il suffit ensuite de renouveler ces différentes étapes sur d'autres sous-bandes et sur les autres calques de la LDI.

5. Conclusion

Nous avons présenté dans cet article le principe d'un schéma de tatouage d'images multi-vues représentées par une structure LDI. Les calques de cette structure appelés calques résiduels ne contiennent que peu d'informations réparties d'une façon très dispersées. L'étape de tatouage a été jointe à la quantification vectorielle algébrique à zone morte qui permet de ne conserver que les vecteurs des coefficients qui correspondent à la description des contours (détails) et d'éliminer les vecteurs des coefficients de valeurs nulles qui correspondent à des zones homogènes. Le schéma de tatouage doit être appliqué à l'ensemble des calques de la LDI avec la répartition des bits du message sur l'ensemble des calques avec un choix adéquat des sous bandes pour chaque calque. Ce travail est encore au stade expérimental. Le développement est en cours de réaliser et les premiers tests devraient être réalisés prochainement.

Références

- [Bat12] BATTIN B. : *Compression multi-vues de flux autostéréoscopiques*. PhD thesis, 2012. Thèse de doctorat dirigée par Lucas, Laurent Informatique Reims 2012.
- [BBD14] BURINI C., BAUDRY S., DOËRR G. : Blind detection for disparity-coherent stereo video watermarking, 2014.
- [BVZ99] BOYKOV Y., VEKSLER O., ZABIH R. : Fast approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. Vol. 23 (1999), 1222–1239.
- [CMB*08a] COX I., MILLER M., BLOOM J., FRIDRICH J., KALKER T. : *Digital Watermarking and Steganography*, 2 ed. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2008.
- [CMB*08b] COX I., MILLER M., BLOOM J., FRIDRICH J., KALKER T. : *Digital Watermarking and Steganography*, 2 ed. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2008.
- [JMG09] JANTET V., MORIN L., GUILLEMOT C. : Incremental-ldi from multi-view coding. In *3DTV-Conference 2009 : The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON 2009)* (2009).
- [KCA10] KOZ A., CIGLA C., ALATAN A. : Watermarking of free-view video. *Image Processing, IEEE Transactions on*. Vol. 19, Num. 7 (July 2010), 1785–1797.
- [KH98] KUNDUR D., HATZINAKOS D. : Digital watermarking using multiresolution wavelet decomposition. In *Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP '98, Seattle, Washington, USA, May 12-15, 1998* (1998), pp. 2969–2972.
- [KZ04] KOLMOGOROV V., ZABIH R. : What energy functions can be minimized via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. Vol. 26 (2004), 147–159.
- [LL00] LI S., LI W. : Shape-adaptive discrete wavelet transforms for arbitrarily shaped visual object coding.

- IEEE Transactions on Circuits and Systems for Video Technology*. Vol. 10, Num. 5 (Aug 2000), 725–743.
- [MCP13] MITREA M., CHAMMEM A., PRÊTEUX F. : Tatouage stéréoscopique. In *Vidéo 3D - Capture, traitement et diffusion*, Lucas L., Loscos C., Remion Y., (Eds.). Hermès - Lavoisier, septembre 2013, pp. 251–272.
- [Niq11] NIQUIN C. : *Reconstruction du relief et mixage réel/virtuel par caméras relief multi-points de vues*. PhD thesis, 2011. Thèse de doctorat dirigée par Remion, Yannick Informatique Reims 2011.
- [NNS72] NEWELL M. E., NEWELL R. G., SANCHAT. L. : A new approach on the shaded picture problem. *Proceedings of the ACM National Conference* (1972), 443 – 450.
- [NPR10a] NIQUIN C., PRÉVOST S., REMION Y. : An occlusion approach with consistency constraint for multi-scopic depth extraction. *International Journal of Digital Multimedia Broadcasting* (2010).
- [NPR10b] NIQUIN C., PRÉVOST S., REMION Y. : A point cloud based pipeline for depth reconstruction from autostereoscopic sets, 2010.
- [Voi03] VOINSON T. : *Quantification vectorielle algébrique avec zone morte : application à la compression d'images à bas débit et au tatouage d'images*. 2003.
- [YLKYS07] YOON S.-U., LEE E.-K., KIM S.-Y., YOSUNG H. : A framework for representation and processing of multi-view video using the concept of layered depth image. *Journal of VLSI Signal Processing*. Vol. 46 (2007), 87–102.