



Ecole Nationale Polytechnique
Département d'Electronique
Laboratoire Signal et Communications



Thèse de Doctorat en Electronique

Option : Signal et Communications

Présenté par :
BOUCHAMA Samira

Magister en Electronique de l'ENP

Intitulé

Les techniques de tatouage appliquées au standard vidéo H.264/AVC

Membres de Jury :

Président :	Mr. Daoud BERKANI,	Professeur,	ENP
Rapporteur :	Mme. Latifa HAMAMI,	Professeur,	ENP
Examineurs :	Mr. Mouloud ADEL,	Maître de Conférences/ HDR,	U. AIX- MARSEILLE
	Melle. Nadja BENBLIDIA	Maître de Conférences/A,	U. S.D.Blida
	Mr. Djamel BOUCHAFFRA,	Directeur de Recherche	CDTA
	Mr. Hichem BOUSBIA-SALAH	Maître de Conférences/ A,	ENP

ENP 2014

ملخص- إن تطبيق تقنيات العلامة المائية على ترميز الفيديو H.264/AVC يتطلب النظر في عدة مواصفات تهدف إلى تلبية تطبيق معين للعلامة المائية و الحفاظ على مستوى الأداء لتشفير الفيديو. وتكمن صعوبة ذلك في تحقيق حل وسط بين المواصفات المختلفة وأيضا في تعقيد الترميز. وقد سمحت لنا دراسة الترميز H.264/AVC و تقنيات العلامة المائية المطبقة على الصور والفيديو باستغلال وحداته لتوفير حلول لإدراج العلامات. نقدم طريقتين « قابلتين للاسترجاع » لإدراج العلامات حيث يمكن استعادة شريط الفيديو الأصلي بعد إدراج العلامة. هاتان الطريقتان تستخدمان معاملات الكم للترميز من أجل إدخال العلامة المائية. تم تطبيق الطريقة الأولى اعتمادا على تقنية إزاحة الرسم البياني المستعملة في الأصل على الصور لرفع نسبة الإدراج. أما الطريقة الثانية فهي مستعملة في الأصل على الصور المضغوطة (JPEG) على مستوى معاملات الكم للترميز. ونقترح أيضا طريقة جديدة تستغل وسائل التنبؤ الخاصة بالترميز لإدراج العلامة المائية. هذا الأسلوب يعطي نتائج أفضل بالمقارنة مع الأساليب المعتمدة على معاملات الكم بشكل عام وخصوصا عندما يتعلق الأمر بقدرة الإدراج والجودة والوقت اللازم للحساب. وأخيرا، فإننا نقدم مشكلة انتشار الخطأ التي تظهر في مرحلة فك التشفير للفيديو المحتوي على علامة مائية. نقترح حلين تم تصميمهما للحد من انتشار هذا الخطأ وبالتالي تحسين جودة الفيديو.

المفاتيح: ترميز الفيديو H.264/AVC، العلامة المائية للفيديو، إدراج قابل للاسترجاع، قدرة الإدراج، جودة الفيديو، الحفاظ على تدفق الفيديو، وضعيات التنبؤ، انتشار الخطأ.

Résumé - Les techniques de tatouage appliquées au codec vidéo H.264/AVC nécessitent la prise en considération de plusieurs critères qui ont pour objectifs de satisfaire une application de tatouage donnée, d'une part, et de préserver les performances du codeur, d'autre part. La difficulté réside dans la réalisation d'un compromis entre les différents critères et dans la complexité du codeur. L'étude du codec H.264/AVC et des techniques de tatouage appliquées à l'image et à la vidéo nous a permis d'exploiter ses modules pour proposer des solutions de tatouage. Nous présentons deux méthodes de tatouage réversible qui permettent de retrouver la vidéo originale après marquage. Ces méthodes utilisent les coefficients DCT quantifiés du codec pour l'insertion de la marque. La première méthode, basée sur le décalage de l'histogramme, a été appliquée initialement aux images pour augmenter la capacité d'insertion. La deuxième méthode a été appliquée initialement aux images compressées (JPEG) et particulièrement sur les blocs de coefficients quantifiés. Nous proposons également une nouvelle méthode qui exploite les modes de prédiction pour l'insertion de la marque. Cette méthode donne de meilleurs résultats par rapport aux méthodes basées sur les coefficients DCT d'une manière générale, notamment quand il s'agit de la capacité d'insertion, de la qualité visuelle et du temps de calcul. Enfin, nous présentons le problème de la propagation de l'erreur qui apparaît dans la phase de décodage du flux vidéo marqué. Nous proposons deux solutions qui ont pour objectif la limitation de la propagation de l'erreur et donc l'amélioration de la qualité visuelle de la vidéo.

Mots clés : Codec H264/AVC, tatouage vidéo, tatouage réversible, capacité d'insertion, qualité de la vidéo, préservation du débit, modes de prédiction, propagation de l'erreur.

Abstract - Watermarking techniques applied to the H.264/AVC video codec need to consider several criteria that aim to satisfy a watermarking application, on one hand, and to preserve the codec performance, on the other hand. The challenge lies in achieving a compromise between the different criteria and also in the codec complexity. The study of the H.264/AVC codec and watermarking techniques applied to the image and video has allowed us to exploit its modules to provide watermarking solutions. We present two reversible watermarking methods that can restore the original video after data embedding. These methods use the DCT quantized coefficients for secret data embedding. The first method, based on the shifted histogram was initially applied to images to increase the embedding capacity. The second method was initially applied to compressed images (JPEG) and especially to the blocks of quantized coefficients. We also propose a new method that exploits the prediction modes to embed the watermark. This method gives better results compared to the DCT- based methods in general, especially when it comes to the embedding capacity, visual quality and computation time. Finally, we present the problem of the error propagation that appears in the decoding phase of the watermarked video stream. We propose two solutions that are designed to limit the spread of the error and thus improve the visual quality of the video.

Keywords : H.264/AVC codec, video watermarking, reversible watermarking, embedding capacity, video quality, bitrate preservation, prediction modes, error propagation.

Remerciements

Je tiens à remercier profondément, à travers cette page, toutes les personnes qui ont contribué à l'accomplissement de ce travail.

Tout d'abord, j'adresse mes sincères remerciements à ma Directrice de Thèse, Professeur Latifa Hamami, pour ses orientations et ses encouragements continus tout au long de mes travaux et aussi pour la confiance qu'elle a eu en moi.

Mes vifs remerciements vont également à tous les membres de jury qui ont bien voulu donner de leur temps précieux pour lire et évaluer ce travail. Le président de jury le Professeur Berkani et les examinateurs : Dr. Adel, Dr. Benblidia, Pr. Bouchaffra et Dr. Bousbia-Salah.

Un grand merci aux professeurs qui m'ont accueilli au sein de leurs laboratoires et qui m'ont accordé une aide inestimable durant les périodes de stage, il s'agit du Professeur Moncef Gabboudj de l'Université des Technologies de Tampere (Finlande), et du Professeur Mohammed Ghanbari de l'Université d'Essex (RU).

Je remercie également tous mes collègues du Centre de Recherche sur l'Information Scientifique et Technique (CERIST) et particulièrement Mme Hassina Aliane pour son aide et encouragements. Merci également à Mmes Houria Zaidi et Karima Zekkar ainsi qu'à toute l'équipe administrative qui a facilité mes départs en stages et en Conférences.

Merci beaucoup à Karima Ait Saadi du Centre de Développement des Technologies Avancées (CDTA), pour la discussion fructueuse que nous avons eu sur le domaine du tatouage de la vidéo.

Merci à tous mes enseignants de l'École Nationale Polytechnique et à mes camarades du Laboratoire Signal et Communications (particulièrement Bahia et Nadia).

Enfin, je voudrais remercier tous mes amis et toute ma famille pour qui je dédie ce travail, particulièrement mes parents, mon frère, mes sœurs et mes neveux. Merci pour l'aide morale que vous m'avez apportée.

Sommaire

Liste des figures
Liste des tableaux
Liste des abréviations

Introduction générale	1
------------------------------------	----------

Chapitre 1: La norme vidéo H.264/AVC.

1.1 Introduction	5
1.2 La compression vidéo : principe de base.....	5
1.2.1 Espace de couleur et échantillonnage	5
1.2.2 Principe de la compression vidéo	6
1.2.3 Les standards de vidéo numérique.....	8
1.3 Le codec H.264/AVC	9
1.3.1 Schéma général.....	9
1.3.2 Hiérarchie des données.....	10
1.3.3 La prédiction.....	12
1.3.3.1 La prédiction Intra	12
1.3.3.2 La prédiction Inter	13
1.3.4 Transformation et quantification	14
1.3.5 Filtre anti-bloc adaptatif	15
1.3.6 Codage entropique.....	16
1.3.6.1 Le codage VLC.....	16
1.3.6.2 Le codage CABAC.....	18
1.3.7 Les profils et niveaux	18
1.4 Conclusion.....	20

Chapitre 2: Les techniques de tatouage vidéo.

2.1 Introduction	21
2.2 Généralités sur les techniques de tatouage vidéo.....	21
2.2.1 Définition du tatouage	21
2.2.2 Applications et critères	22
2.2.3 Quelques techniques de tatouage.....	23
2.2.4 Domaine d'insertion	24
2.2.5 Les techniques de tatouage appliquées à la vidéo.....	24
2.2.6 Les attaques	25
2.3 Les techniques de tatouage appliquées au standard H.264/AVC	26
2.3.1 Les exigences du tatouage vidéo	26
2.3.2 La zone d'insertion	26
2.3.2.1 Insertion de la marque au cours du processus de compression.....	27
2.3.2.2 Insertion de la marque après compression.....	29
2.3.3 Schémas de tatouage réversible	30
2.3.4 Le prétraitement de la marque	31
2.4 Evaluation de la qualité vidéo	32
2.4.1 Notion d'artéfacts	32
2.4.2 Evaluation subjective de la qualité	33
2.4.3 Les métriques de qualité	34
2.5 Conclusion.....	36

Chapitre 3: Méthodes de tatouage réversible.

3.1 Introduction	37
3.2 Le tatouage réversible.....	37
3.3 Méthode réversible basée sur le décalage de l'histogramme.....	38
3.3.1 La méthode de Ni	38

3.3.2 Approche proposée	39
3.3.2.1 Insertion et extraction de la marque.....	39
3.3.2.2 Evaluation de la qualité de la vidéo.....	40
3.3.2.3 Résultats et analyses	41
3.4 Méthode réversible basée sur les séquences nulles des blocs DCT.....	46
3.4.1 Les méthodes de Cheng et Gujjunoori.....	46
3.4.2 Approche proposée	47
3.4.2.1 Schéma adapté	47
3.4.2.2 Nouveau schéma.....	48
3.4.2.3 Résultats et analyses	49
3.5 Conclusion	53

Chapitre 4: Méthode de tatouage basée sur la prédiction.

4.1 Introduction	54
4.2 La prédiction Intra-images.....	54
4.3 Les modes de prédiction intra 4x4.....	55
4.4 Le mode de décision optimal pour la prédiction intra 4x4	55
4.5 Le mode le plus probable.....	56
4.6 Méthode de marquage basée sur la prédiction intra et les règles de mise en correspondance.....	56
4.7 L'approche proposée	58
4.7.1 Présentation générale.....	58
4.7.2 Processus d'insertion	59
4.7.3 Processus d'extraction	61
4.7.4 La sécurité et la robustesse du tatouage.....	61
4.8 Résultats et analyses	62
4.9 Conclusion	66

Chapitre 5: Compensation de la propagation d'erreur.

5.1 Introduction	67
5.2 La propagation de l'erreur	68
5.2.1 Les processus de codage et de décodage du codec H.264/AVC.....	68
5.2.2 Description de la propagation de la distorsion.....	69
5.2.3 L'estimation de l'erreur	70
5.3 Limitation de la propagation d'erreur basée sur le marquage de paires de coefficients.....	71
5.3.1 Elimination de la propagation d'erreur.....	71
5.3.2 Approche proposée: Premier schéma	72
5.4. Limitation de la propagation d'erreur basée sur la compensation d'erreur dans les blocs voisins.....	74
5.4.1 La compensation de l'erreur	74
5.4.2 Approche proposée : Deuxième schéma.....	75
5.5 Résultats et analyses	76
5.5.1 Premier schéma.....	76
5.5.2 Deuxième schéma.....	78
5.5.3 Comparaison avec d'autres méthodes	80
5.6 Conclusion	81

Conclusions et perspectives.....83

Bibliographie	5
Annexe 1: Contributions de l'auteur.....	10
Annexe 2: Glossaire francais-anglais	11

Liste des figures

Figure 1 : Schéma général d'un codec vidéo	2
Figure 2 : Schéma général de marquage du codec H.264/AVC.....	2
Figure 1.1 : Echantillonnage 4:2:0, 4:2:2 et 4:4:4	6
Figure 1.2 : Schéma représentant les images I, P et B dans une séquence vidéo.....	7
Figure 1.3 : Un macrobloc de taille 16x16 dans un bloc de recherche d'un paramètre de 7 pixels.....	7
Figure 1.4 : Espace de recherche de la prédiction de mouvement.....	7
Figure 1.5 : Les standards ISO et ITU.....	9
Figure 1.6 : Schéma simplifié du codage H.264/AVC.....	10
Figure 1.7 : Schéma simplifié du décodage H.264/AVC.....	10
Figure 1.8 : Hiérarchie des données dans le flux vidéo.....	11
Figure 1.9 : Séquence vidéo entrelacée	12
Figure 1.10 : Les modes de prédiction intra 16x16.....	12
Figure 1.11 : Partitions (16x16, 8x16, 16x8, 8x8) et sous partitions (8x8, 4x8, 8x4, 4x4) du macrobloc.....	13
Figure 1.12 : Exemples de vecteurs de mouvements.....	14
Figure 1.13 : Ordre de transmission des blocs résiduels dans le mode Intra 16x16.....	15
Figure 1.14 : Illustration des bordures horizontales et verticales.....	16
Figure 1.15 : (a) Exemple de bloc 4x4, (b) Lecture en zigzag des coefficients.....	17
Figure 1.16 : Profils de la norme H.264/AVC.....	19
Figure 2.1 : Insertion de la marque basée sur la méthode de l'étalement du spectre.....	23
Figure 2.2 : Classification des méthodes de tatouage vidéo.....	24
Figure 2.3 : Les attaques du tatouage.....	25
Figure 2.4 : (a) Le balayage en Zig-zag, (b) Le balayage en champs, (c) Exemple de zones d'insertion dans un bloc de coefficients DCT.....	28
Figure 2.5 : Exemples d'artefacts : (a) Image originale de Barbara (b) L'effet de bloc, (c) Le flou.....	32
Figure 2.6 : Structure de test des séquences vidéo.....	34
Figure 3.1 : (a) L'image de Barbara (b) L'histogramme de Barbara.....	39
Figure 3.2 : Forme de l'histogramme des coefficients DCT quantifiés pour la séquence Coast-Guard.....	40
Figure 3.3 : Histogramme des coefficients DCT significatifs de l'image I de la séquence Walk : (a) Avant décalage, (b) Après décalage.....	41
Figure 3.4 : Séquence Coast-Guard (a) Image I non marquée, (b) Image I marquée, (c) Image P marquée.....	42
Figure 3.5 : (a1), (a2) images I des séquences non marquées Walk et Silent respectivement, (b1) et (b2) sont les images marquées correspondantes.....	42
Figure 3.6 : Séquence Foreman : (a) Image I non marquée, (b) Image I marquée, (c) Image P marquée.....	42
Figure 3.7 : PSNR des séquences vidéo marquées et non marquées: (a) Walk, (b) Coast-Guard, (c) Silent, et (d) Foreman.....	43
Figure 3.8 : (a) Ensembles d'insertion (b) Exemple de coefficients quantifiés.....	46
Figure 3.9 : (a) Ensemble R pour l'insertion des données.....	48
Figure 3.10 : (a1), (a2) et (a3) représentent les images I non marquées des séquences Coast-Guard, Hall et Mobile respectivement. (b1), (b2) et (b3) sont les images correspondantes marquées. (c1), (c2) et (c3) sont les premières images P correspondantes marquées.....	50
Figure 3.11 : Capacité d'insertion pour six séquences vidéo pour trois méthodes.....	52
Figure 3.12 : Séquence Mother, 13 ^{ème} image: (a) Séquence non marquée, (b) Séquence marquée selon Gujjunoori, (c) Séquence marquée selon notre méthode.....	52
Figure 3.13 : Séquence Mobile, 25 ^{ème} image: (a) Séquence non marquée, (b) Séquence marquée selon notre méthode (SH), (c) Séquence marquée selon notre méthode (DCT).....	52
Figure 4.1 : (a) Blocs voisins se positionnant en haut et à gauche du bloc en cours, (b) Echantillons de prédiction.....	55
Figure 4.2 : Les neuf directions pour la prédiction Intra.....	55

Figure 4.3 : Histogramme des modes de prédiction Intra 4x4 d'une image I de la séquence Bridge-Close: (a) Etat initial, (b) Apres création du vide dans le mode DC.....	59
Figure 4.4: Image I de la séquence Container : Macroblocs codés en Intra 4x4 et en Intra 16x16.....	62
Figure 4.5 : Capacité d'insertion pour cinq séquences tatouées avec notre méthode et avec les méthodes proposées par Yang, Hu et Xu.....	63
Figure 4.6 : Augmentation du nombre de macroblocs codés en intra 16x16 dans la première image I de la séquence Silent: (a) la séquence non marquée, (b) la séquence marquée.....	64
Figure 4.7 : (a1) et (a2) représentent les images I non marquées des séquences Bridge et Container, et (b1) et (b2) sont les images marquées.....	64
Figure 4.8 : (a) et (b) représentent le temps d'encodage et de décodage des séquences de tests marquées et non marquées.....	65
Figure 5.1 : (a) et (b) représentent un schéma général des processus.....	69
Figure 5.2 : (a) Mode de Prédiction n° 3, (b) Blocs et échantillons de prédiction.....	69
Figure 5.3 : Blocs voisins se positionnant à droite et en bas du bloc en cours.....	72
Figure 5.4 : (a) Représentation des positions des paires de coefficients, (b), (c), (d), (e), (f) et (g) L'estimation de l'erreur introduite dans les échantillons de bloc en insérant un bit dans (P_1, P_8) , (P_2, P_7) , (P_3, P_{11}) , (P_5, P_{11}) , (P_6, P_{13}) et (P_9, P_{14}) respectivement.....	73
Figure 5.5 : Erreur propagée au bloc A via le mode 3.....	75
Figure 5.6 : Les différences de PSNR calculées pour 8 séquences pour $N>5$, $N>7$ et $N>9$	77
Figure 5.7 : La séquence Hall, image P n° 33 (a) image non marquée, (b) image marquée ($N>5$), (c) image marquée ($N>7$), (d) image marquée ($N>9$).....	77
Figure 5.8 : Le débit de 8 séquences marquées ou non pour $N>5$, $N>7$ et $N>9$	78
Figure 5.9 : Le débit de 8 séquences marquées ou non pour $N>5$, $N>7$ et $N>9$	78
Figure 5.10: Différences de PSNR mesurées pour huit séquences.....	79
Figure 5.11: Séquence Mother, (a) Image non marquée, (b) Image marquée sans correction d'erreur, (c) Image marquée avec correction d'erreur.....	79

Liste des tableaux

Tableau 1.1. Mots –codes Golomb exponentiel.	17
Tableau 3.1. La capacité d’insertion, l’augmentation de débit et la qualité vidéo de la séquence Walk.....	44
Tableau 3.2. La capacité d’insertion, l’augmentation de débit et la qualité vidéo de la séquence CG.....	44
Tableau 3.3. La capacité d’insertion, l’augmentation de débit et la qualité vidéo de la séquence Silent.....	44
Tableau 3.4. La capacité d’insertion, l’augmentation de débit et la qualité de la vidéo de la séquence Foreman. ...	45
Tableau 3.5. Règles de mise en correspondance pour l’insertion.....	49
Tableau 3.6. Capacité d’insertion, PSNR et augmentation du débit pour six séquences vidéo.....	51
Tableau 3.7. Comparaison entre nos approches DCT et SH et celle de Gujjunoori.....	51
Tableau 4.1. Règles de correspondance liées à l’insertion des bits 00 dans un groupe de bloc.	57
Tableau 4.2. Configuration des modes pour l’insertion de la marque pour les groupes 1 et 2.....	59
Tableau 4.3. Configuration des modes pour l’insertion de la marque pour les groupes 3 et 4.....	60
Tableau 4.4. Les résultats obtenus pour huit séquences vidéo.	62
Tableau 4.5. Encodage des premières images de la séquence Silent non marquée et marquée.	65
Tableau 5.1. Exemple de matrice d’erreur dans le cas d’insertion d’un bit dans un coefficient ou une paire de coefficients.....	72
Tableau 5.2. Processus d’insertion	74
Tableau 5.3. Configuration du H.264/AVC.	76
Tableau 5.4. Augmentation du débit et de la capacité l’insertion pour huit séquences vidéo.....	79
Tableau 5.5. Comparaison entre nos approches et celle de Ma.....	80
Tableau 5.6. Comparaison entre nos approches et celle de Huo	81

Liste des abréviations

AC	Alternative Current
ASO	Arbitrary Slice Order
AVC	Advanced Video Coding
B	Bi-predictive frame
Bs	Boundary-strength
CABAC	Context-Adaptive Binary Arithmetic Coding
CAVLC	Context-Adaptive Variable Length Coding
Codec	Codeur/ decodeur.
DC	Direct Current
DCT	Discret Cosine Transform
DE	Difference Expansion
DRM	Digital Right Management
DVD	Digital Versatile Disc
FRExt	Fidelity Range Extensions
FMO	Flexible Macroblock Ordering
GBSD	Generic Syntax Description Bitstream
GOP	Group Of Picture
I	Intra predictive frame
IEC	International Electrotechnical Commission
IPCM	Intra Pulse Code Modulation
ISO	International Standardization Organisation
ITU-T	International Telecommunication Union-Telecommunication
JPEG	Joint Photographic Experts Group
JVT	Joint Video Team
JTC	Joint Technical Committee
LSB	Least Significant Bit
MBAFF	Macroblock-Adaptive Frame-Field
MPEG	Moving Pictures Expert Group
MPM	Most Probable Mode
MSE	Mean Square Error
NAL	Network Adaptation Layer
P	Predictive frame
PAFF	Picture-Adaptive Frame-Field
POC	Picture Order Count
PSNR	Peak Signal to Noise Ratio
RCM	Reversible Contrast Mapping
RDO	Rate Distortion Optimization
RDC	Rate Distortion Cost
RS	Redundant Slices
RVB ou RGB	Rouge-Vert-Bleu (ou Red-Green-Blue)
SH	Shifted Histogram
SI	Switched I
SP	Switched P
VCL	Video Coding Layer
VLC	Variable Length Coding
VCEG	Video Coding Experts Group
YCbCr	Symbole de luminance (Y), chrominance (Cb) et saturation (Cr)

Introduction générale

L'intérêt de la vidéo est aujourd'hui incontestable pour différents domaines tels que la médecine, le marketing, l'enseignement et le divertissement. La production et l'exploitation de la vidéo numérique sont devenues à la portée de toutes les tranches d'âge. Elles sont favorisées par la baisse des coûts du matériel et logiciel de production de la vidéo, ainsi que l'adoption massive d'internet dans les foyers et les entreprises.

Par ailleurs, le développement des outils de compression a permis la réduction du coût du transfert de la vidéo et de son archivage. De plus, la mise en place des serveurs de vidéo et le développement de la technique de streaming pour la diffusion de la vidéo à travers des sites tels que Youtube¹, Dailymotion², Hulu³, etc., a sans doute encouragé la production, la consommation, le partage, et la rediffusion de la vidéo.

Toutefois, de cette situation sont créés de nouveaux problèmes liés notamment à la protection de la propriété intellectuelle et à l'authentification du document. Le besoin de protéger la vidéo après sa réception se fait ressentir. Les techniques classiques de protection de droits d'auteurs ne s'adaptent plus à l'ère du numérique et les techniques de cryptographie restent insuffisantes [1].

L'introduction des techniques de tatouage a été donc proposée initialement pour la protection des droits d'auteurs, ensuite pour d'autres applications telles que l'authentification, le contrôle d'intégrité et la confidentialité des données. Ceci se fait par l'insertion de données appelées marque ou signature dans l'œuvre multimédia pour attester de l'identité du propriétaire du document.

Pour la vidéo, les chercheurs tentent d'adapter les méthodes de tatouage existantes ou de concevoir de nouvelles méthodes spécifiques à ce support. Le codec H.264/AVC est particulièrement concerné par le tatouage, toutefois, le taux de compression élevé qu'il offre constitue un défi technique pour l'insertion de la marque.

La norme vidéo H.264/AVC est le résultat de travail entre le groupe d'experts en codage vidéo VCEG (de l'ITU-T⁴) et le groupe d'experts d'images animées MPEG (de l'ISO/IEC⁵). Elle cible un large éventail d'applications à débits variables telles que la télévision, la vidéo conférence, le streaming vidéo, le stockage sur disque ou DVD, etc. Elle est appelée à être plus adoptée que les normes précédentes étant donné l'amélioration considérable qu'elle apporte [2][3].

¹ www.youtube.com

² www.dailymotion.com

³ www.hulu.com

⁴ International Telecommunication Union- Telecommunication.

⁵ International Standardization Organisation/ International Electrotechnical Commission.

Plusieurs méthodes de tatouage vidéo ont été proposées pour ce codec, promettant de répondre à certains critères tels que l'imperceptibilité, la grande capacité d'insertion ou encore le maintien du débit, pour satisfaire une application donnée. La difficulté réside dans la réalisation d'un compromis entre les différents critères.

Le codec H.264/AVC est composé essentiellement de trois modules exploités pour l'insertion de la marque, à savoir : La prédiction, la transformation et la quantification, et le codage entropique. Ces modules sont représentés sur la figure 1.

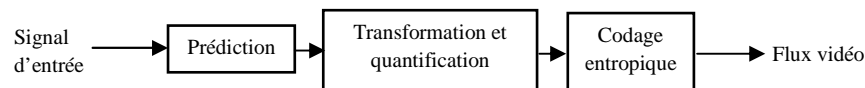


Figure 1 : Schéma général d'un codec vidéo [3].

L'insertion de la marque peut être effectuée après l'étape de compression en introduisant un décodeur entropique avant d'effectuer l'insertion, ou directement durant le processus de compression, ce qui serait probablement plus adapté pour les applications en temps réel. Un schéma général du marquage du codec H.264/AVC est illustré sur la figure 2. Les flèches en pointillés indiquent les éventuelles zones d'insertion de la marque.

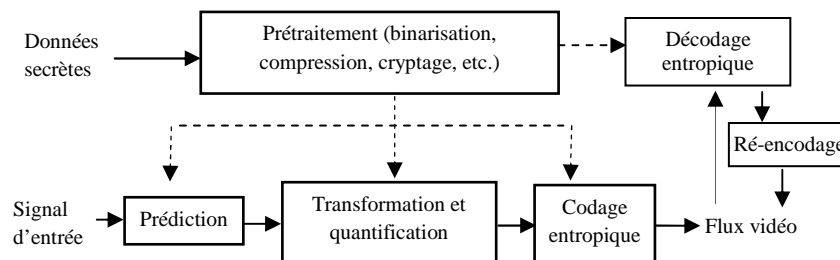


Figure 2 : Schéma général de marquage du codec H.264/AVC.

Dans le cadre de notre travail, nous nous intéressons aux techniques de tatouage appliquées au codec H.264/AVC en explorant les deux premiers modules pour l'insertion de la marque durant le processus de compression.

Dans la littérature, les premiers travaux de marquage ont ciblé les coefficients quantifiés obtenus après l'étape de transformation et de quantification [4][5]. Les résultats sont généralement présentés par trois critères : la capacité d'insertion qui est un nombre N de bits insérés dans une image de la vidéo (ou dans toute la séquence), l'augmentation du débit et la qualité visuelle souvent obtenue par une métrique objective. Le défi réside dans la possibilité d'augmenter la capacité d'insertion N tout en préservant (ou en améliorant) les résultats par rapport au débit et à la qualité visuelle de la vidéo. Toutefois, très peu de méthodes satisfont le critère de réversibilité, souvent utile, notamment dans le cadre de certaines applications médicales ou militaires. Ce critère permet de récupérer à partir de la vidéo tatouée, la vidéo originale, si l'utilisateur final le

souhaite. Dans le cadre de ce travail, une première contribution serait de satisfaire le critère de réversibilité en adaptant à la vidéo compressée des méthodes de tatouage réversible qui ont été proposées initialement pour les images. Evidemment, l'adaptation devrait tenir compte également des critères cités précédemment.

D'autres méthodes plus récentes exploitent les modes de prédiction pour l'insertion de la marque [6][7][8], ces méthodes donnent des résultats nettement meilleurs par rapport au compromis entre les trois critères cités précédemment (capacité d'insertion, débit, qualité visuelle). Dans un deuxième travail nous nous intéressons à ce type de méthodes en proposant une nouvelle approche qui semblerait bien adaptée pour les applications en temps réel qui nécessitent relativement de grandes capacités d'insertion.

Toutefois, les méthodes basées sur les coefficients quantifiés restent plus attractives aux chercheurs car, jusqu'à présent, elles peuvent satisfaire beaucoup plus de critères, telle la réversibilité, par rapport aux méthodes basées sur les modes de prédiction. C'est la raison pour laquelle nous revenons aux méthodes basées sur les coefficients quantifiés et nous nous penchons sur le problème de la propagation de la distorsion qui apparaît au niveau du décodeur H.264/AVC. Ce problème se traduit par la propagation de l'erreur introduite par le tatouage d'un bloc d'une image vidéo à toute la séquence vidéo, contraignant les auteurs à limiter les zones d'insertion et par conséquent la capacité. De ce fait, une solution à ce problème constituerait sans doute un élément clé pour améliorer les performances d'un système de tatouage, notamment s'il est basé sur l'exploitation des coefficients quantifiés.

Ce document est organisé comme suit :

Le premier chapitre donne une brève description de la compression vidéo et des différents modules du codec vidéo H.264/AVC particulièrement. Dans le second chapitre nous présentons des généralités sur les techniques de tatouage appliquées à la vidéo et particulièrement un état de l'art sur les techniques de tatouage appliquées au codec H.264/AVC.

Dans le troisième chapitre nous présentons deux méthodes de tatouage réversible que nous introduisons au codec H.264/AVC en se basant sur les coefficients de la transformée en cosinus discrète (DCT) quantifiés. La première méthode, basée sur le décalage de l'histogramme, a été appliquée initialement aux images pour augmenter la capacité d'insertion. La deuxième méthode a été appliquée initialement aux images JPEG et particulièrement sur les blocs de coefficients quantifiés de taille 8x8, une adaptation est nécessaire pour l'application de cette méthode sur les blocs de coefficients quantifiés de taille 4x4 qu'on retrouve au niveau du codec vidéo.

Dans le quatrième chapitre nous proposons une méthode originale qui exploite les modes de prédiction pour l'insertion de la marque. Cette méthode donne de meilleurs résultats par rapport aux méthodes basées sur la DCT d'une manière générale, notamment quand il s'agit de la capacité d'insertion, la qualité visuelle et le temps de calcul.

Dans le cinquième chapitre, nous présentons le problème de la propagation de l'erreur et nous proposons deux solutions nouvelles qui ont pour objectif la limitation de la propagation et donc l'amélioration de la qualité visuelle de la vidéo. Dans la première solution, nous effectuons une adaptation de la méthode proposée par Ma et al. [9] pour une insertion durant le processus de compression. Dans la deuxième solution nous proposons une nouvelle approche pour une compensation de l'erreur propagée en effectuant une correction des blocs qui sont voisins au bloc tatoué.

Enfin, nous présentons dans la conclusion une synthèse des résultats obtenus et les éventuelles perspectives de ce travail.

1.1 Introduction

Avec le développement de la technologie et le passage au numérique, un nouveau défi est apparu afin de compresser la vidéo pour diminuer sa taille de stockage et rendre possible son transport sur des réseaux de communications. Plusieurs normes de codage vidéo sont apparues durant ces dernières décennies, apportant les descriptions et les recommandations pour le codage et le décodage de la vidéo numérique.

La norme vidéo MPEG 4 Part 10 Advanced Video Coding (AVC), connue également sous le nom de H.264, est le résultat de travail entre le groupe d'experts en codage vidéo (VCEG) de l'ITU-T et le groupe d'experts d'images animées (MPEG) de l'ISO/IEC.

Finalisée en mars 2003, cette norme est appelée à être plus adoptée que les normes précédentes étant donné l'amélioration considérable qu'elle apporte. Elle permet de cibler une large gamme d'applications nécessitant des débits variables telles que la télévision, la vidéo conférence, le streaming vidéo, le stockage sur disque ou DVD, etc.

Par rapport aux standards précédents, de nouveaux outils ont été introduits aux différents modules du codage permettant ainsi d'augmenter significativement l'efficacité du codage, telles la prédiction spatiale directionnelle pour le mode Intra, la compensation de mouvement sur des blocs de tailles variables avec une précision d'un quart de pixel, les images de références multiples, la transformée en cosinus discrète (DCT) entière 4x4, en plus de l'utilisation du codage arithmétique adaptatif contextuel [2].

Dans ce chapitre sont présentés la structure générale et les composants d'un système de codage vidéo standard. La grande partie est réservée à la norme H.264/AVC.

1.2 La compression vidéo : principe de base

1.2.1 Espace de couleur et échantillonnage

L'espace de couleur indique la méthode utilisée pour représenter l'information de luminance (ou luma) et la couleur. Les deux méthodes les plus utilisées sont RVB (ou RGB) et YCbCr. Dans l'espace de couleurs RVB, l'image est séparée en trois composants qui représentent respectivement la quantité de rouge (R), de vert (V) et de bleu (B). Par contre, dans l'espace de couleurs YCbCr, les trois composants représentent respectivement la luminance Y, et les composantes de chrominance (ou chroma) Cb et Cr, qui sont la différence entre le B ou le R et la luminance [3][10].

Les équations de conversion d'un système à l'autre sont données par :

$$\begin{aligned} Y &= 0.299 R + 0.587 G + 0.114 B \\ Cb &= 0.564 (B - Y) \\ Cr &= 0.713 (R - Y) \end{aligned} \quad (1)$$

$$\begin{aligned} R &= Y + 1.402 Cr \\ G &= Y - 0.344 Cb - 0.714 Cr \\ B &= Y + 1.772 Cb \end{aligned} \quad (2)$$

L'espace YCbCr s'avère plus efficace, étant donné que la perception humaine est plus sensible aux variations de la luminance que de la chrominance, la quantité d'information de chrominance peut être réduite par rapport à la luminance lorsqu'on utilise l'espace colorimétrique YCbCr. Quand il y a la même quantité d'échantillons Y, Cb et Cr, on dit que l'image est dans le format 4:4:4 et a une haute fidélité de couleurs.

L'échantillonnage 4:2:2 signifie deux échantillons Cb et deux Cr pour chaque quatre échantillons Y. L'échantillonnage 4:2:0 signifie un échantillon Cb et un Cr pour chaque quatre échantillons Y, tel qu'illustré à la figure 1.1.

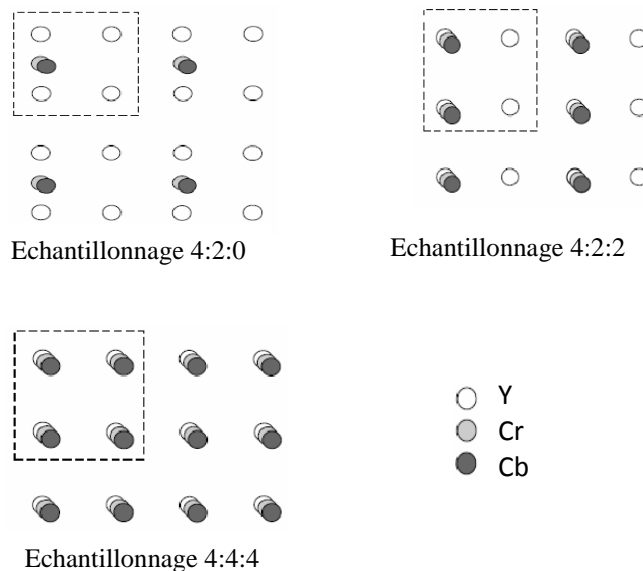


Figure 1.1 : Echantillonnage 4:2:0, 4:2:2 et 4:4:4 [3].

1.2.2 Principe de la compression vidéo

Les systèmes de compression vidéo exploitent les redondances spatiales (dans une même image) et temporelles (ressemblance probable entre les images voisines), et les éliminent pour réduire la quantité et le débit d'informations à stocker ou à transmettre.

Un algorithme de compression commence par encoder la première image d'une scène avec un algorithme de compression d'image fixe. Il la compare ensuite avec la suivante, et n'encode qu'une représentation de leur différence.

Il est possible de retrouver au sein d'une séquence vidéo compressée : des images codées en Intra (I) qui sont des images "références" à partir desquelles est réalisé le décodage. Chaque

changement de plan doit commencer par une image I, ce qui permet également un accès aléatoire en lecture. Les images Prédites (P) se basent sur les technologies liées à l'estimation et à la compensation des mouvements. Elles se basent sur les images codées antérieurement et sont plus compressées que les images I. Les images Bidirectionnelles (B) sont également des images prédites, mais elles présentent la particularité de pouvoir être interpolées à partir d'images I ou P passées et /ou futures comme illustré à la figure 1.2. Les images B ne propagent pas les erreurs puisqu'elles ne sont jamais utilisées comme références.

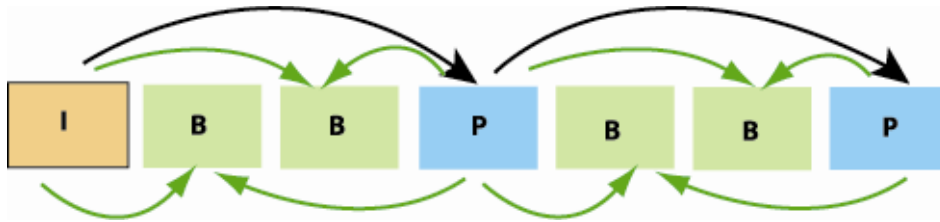


Figure 1.2 : Schéma représentant les images I, P et B dans une séquence vidéo.

Pour l'estimation et la compensation de mouvement, les méthodes de mise en correspondance ou encore "méthodes par appariement de bloc" sont les plus utilisées par les standards vidéo. Elles consistent à faire la correspondance entre des blocs d'images consécutives pour trouver le vecteur qui apparie le mieux une région. Ces méthodes sont basées sur la division de l'image en blocs, chacun d'eux est comparé au bloc correspondant et son voisinage dans l'image précédente, afin de créer des vecteurs de mouvement qui représentent le mouvement d'une image à une autre. La recherche se fait dans une surface de carré selon un paramètre p , tel qu'illustré à la figure 1.3 et à la figure 1.4 [11].

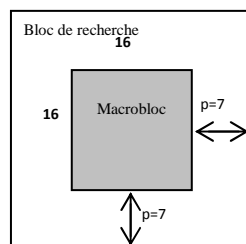


Figure 1.3 : Un macrobloc de taille 16x16 dans un bloc de recherche d'un paramètre de 7 pixels.

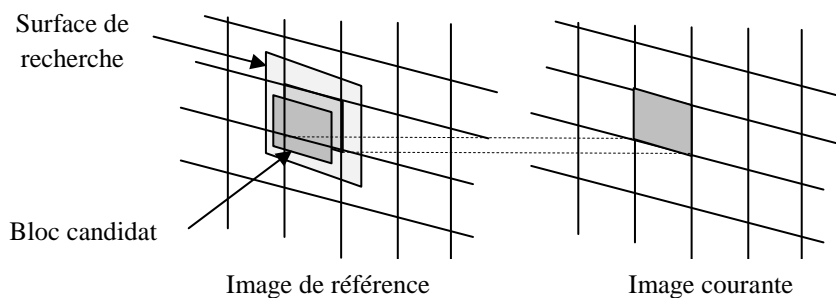


Figure 1.4 : Espace de recherche de la prédiction de mouvement.

Après l'étape de prédiction citée précédemment, l'étape de transformation a pour rôle de mettre en forme le signal d'erreur (la différence entre le signal prédit et le signal original) pour le quantifier et éliminer les redondances des reliquats (données résiduelles ou résidus) pour réduire le débit. Le codage entropique permet de réduire d'avantage les redondances des coefficients ainsi obtenus. Pour plus de détails sur ces différentes étapes, le lecteur peut se référer à [3] et [4].

Les étapes de codage vidéo peuvent être résumées dans les points suivants [12] :

- la prédiction entre les images est assurée par la technique de l'estimation de mouvement qui permet de réduire les redondances temporelles ;
- la prédiction intra-image permet de réduire les redondances ou les corrélations spatiales ;
- le calcul des données résiduelles résultant de la différence entre l'image prédite et l'image source en cours de codage ;
- l'utilisation d'une transformée spatiale discrète, telle la DCT, et d'un filtrage pour éliminer les redondances spatiales des résidus ;
- la quantification permet de réduire la précision (et donc le débit) des coefficients résultants de la transformée spatiale des résidus;
- le codage entropique des coefficients et des données associées telles que les vecteurs de mouvement inter-image.

1.2.3 Les standards de vidéo numérique

Les normes de vidéo numérique ont évolué sous les deux noms : H.26x et MPEG-x. Les codecs H.26x sont recommandés par le secteur de normalisation des télécommunications ITU-T. Les recommandations de l'ITU-T ont été conçues pour les applications de télécommunications telles que la vidéoconférence et la visiophonie. Les produits MPEG-x sont les travaux du comité technique mixte numéro 1 (JTC 1). Les normes MPEG ont été conçues essentiellement pour répondre aux besoins de stockage vidéo (par exemple, CD-ROM, DVD), de diffusion TV et de la vidéo en streaming [4].

La figure 1.5 représente les standards de codage vidéo qui se sont succédés jusqu'à l'apparition du H.264/AVC. Cette figure montre qu'il y a eu des travaux conjoints entre les deux organisations qui ont commencé dès l'année 1984 pour aboutir à la définition du MPEG-2/H.262. Les deux groupes d'experts VCEG et MPEG ont à nouveau mis leurs travaux en commun au sein de l'équipe JVT (*Joint Video Team*) pour aboutir à une nouvelle norme adoptée par les deux organismes sous un nom différent : ITU-T H.264 et ISO MPEG-4 Part 10/AVC [12].

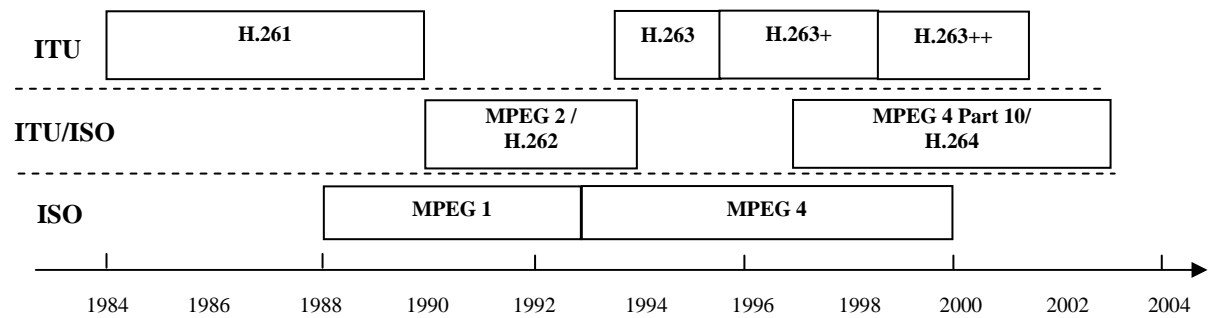


Figure 1.5 : Les standards ISO et ITU.

1.3 Le codec H.264/AVC

Dans cette section, nous présentons des généralités sur les différents modules du codec H.264/AVC. Plus de détails peuvent être retrouvés dans les références [2][3][13].

1.3.1 Schéma général

L'image est d'abord divisée en blocs. La première image d'une séquence est généralement codée en mode "intra". Chaque bloc dans une image intra est prédit en utilisant des pixels voisins des blocs préalablement codés. Les autres images d'une séquence sont généralement codées en "inter" en se basant sur les techniques d'estimation et de compensation de mouvement à partir des images précédemment décodées (images de référence). Les données résiduelles de la prédiction qui représentent la différence entre le bloc original et celui prédit, vont subir une transformation, une quantification, ensuite un codage entropique. D'autres informations nécessitent également un codage entropique, telles que celles relatives au mode de prédiction des macroblocs, aux paramètres de quantification, à la prédiction Inter, etc.

Pour obtenir les blocs des images précédemment codés dans le but de les utiliser comme données de référence, le codeur contient un décodeur qui applique une quantification et une transformation inverses à la sortie du quantificateur tel qu'indiqué dans le schéma général (figure 1.6).

Un filtre anti-bloc est également appliqué sur les bordures horizontales et verticales des blocs 4x4 pour supprimer les artefacts (voir chapitre 2, section 2.4) causés par le processus d'encodage.

Chaque macrobloc dans une image est prédit à partir de données déjà encodées et reconstruites; on garantit ainsi l'utilisation de la même prédiction que le décodeur. Le codeur d'entropie va recevoir la différence (le résidu) entre le bloc courant et la prédiction. Les macroblocs I sont prédits en utilisant les blocs de l'image courante. Les macroblocs P sont prédits en utilisant une à plusieurs images antérieures. Les macroblocs B quant à eux, utilisent des images de références

multiples. Ils sont donc prédits à partir d'une ou plusieurs images, futures et/ou antérieures comme indiqué précédemment.

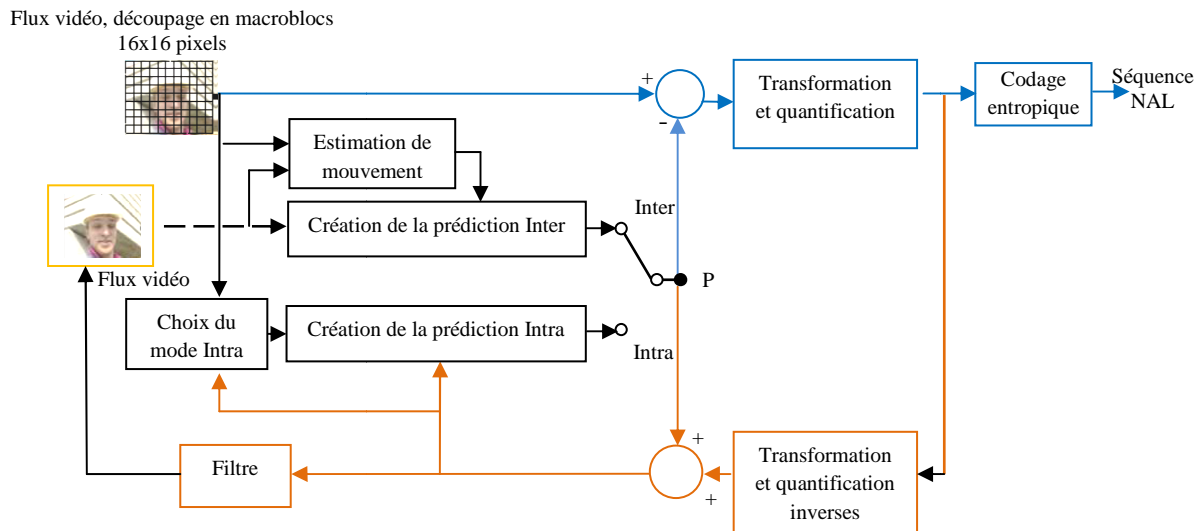


Figure 1.6: Schéma simplifié du codage H.264/AVC.

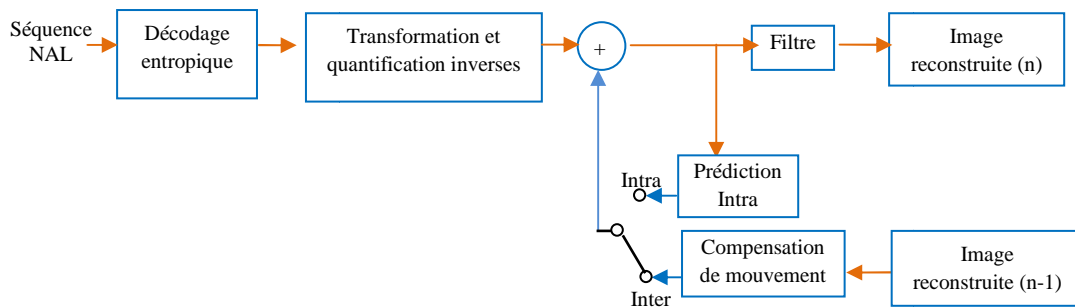


Figure 1.7 : Schéma simplifié du décodage H.264/AVC.

La norme H264/AVC est divisée en deux couches : une première couche de codage vidéo (VCL - Video Coding Layer) s'occupe de représenter efficacement le contenu vidéo, et la couche d'adaptation au réseau (NAL - Network Adaptation Layer) qui s'intéresse plus particulièrement à une mise en forme des données vidéo adaptée au support de transmission.

Le décodeur reçoit le flux compressé et lui applique un décodage entropique. Les coefficients quantifiés sont récupérés pour subir une quantification et une transformation inverses. Le décodeur crée également les blocs prédits à partir des entêtes transmis dans le flux. Les blocs prédits sont rajoutés aux blocs de résidus, ensuite filtrés pour obtenir les blocs reconstruits. Un schéma simplifié du décodeur est représenté à la figure 1.7.

1.3.2 Hiérarchie des données

Une séquence vidéo est une succession d'images, dont chacune contient un numéro (POC) nécessaire pour son réordonnement car les processus d'encodage et de décodage peuvent les

traiter dans le désordre. Les images encodées précédemment (images de référence) sont décodées et organisées dans une ou deux listes pour être utilisées dans la prédiction.

Une image est constituée d'un nombre de macroblocs (des blocs de taille 16×16 pixels), chacun d'eux contenant 16×16 luma (Y) et deux 8×8 chroma (Cb et Cr). Chaque macrobloc peut être décomposé en blocs de 8×8 pixels ou bien 4×4 pixels.

Dans chaque image, les macroblocs sont ordonnés en tranches de gauche à droite et de haut en bas. La longueur d'une tranche varie de un à tous les macroblocs de l'image.

Une image peut contenir de une à plusieurs tranches de différents types (I, P, B, SP, SI). Les tranches I contiennent des macroblocs I, les tranches B contiennent des macroblocs B et I et les tranches P contiennent des macroblocs P et I. SP et SI (le S pour *Switching*) sont des tranches de commutation, qui ont un codage spécifique permettant la permutation efficace entre les flux vidéo lors du décodage.

Les images sont regroupées en des ensembles appelés GOP afin de contourner le problème d'accès aléatoire aux images. En effet, l'accès à une image inter-prédite dans une séquence vidéo pourrait nécessiter le décodage de toutes les images codées préalablement. La figure 1.8 représente la hiérarchie des données dans le flux vidéo.

Une image vidéo peut être considérée comme deux champs intercalés, un champ supérieur et un champ inférieur (voir figure 1.9). De ce fait, il existe deux modes de codage : le mode codage-image qui permet de coder l'image comme étant une seule unité, et le mode codage-champs qui permet de coder chaque champ séparément. Ce dernier mode est efficace lorsque le premier champ est codé en utilisant les tranches I et le second est déduit du premier en utilisant la compensation de mouvement. Le codeur H.264 permet la commutation adaptative entre ces deux modes dans la même image pour assurer plus d'efficacité [2].

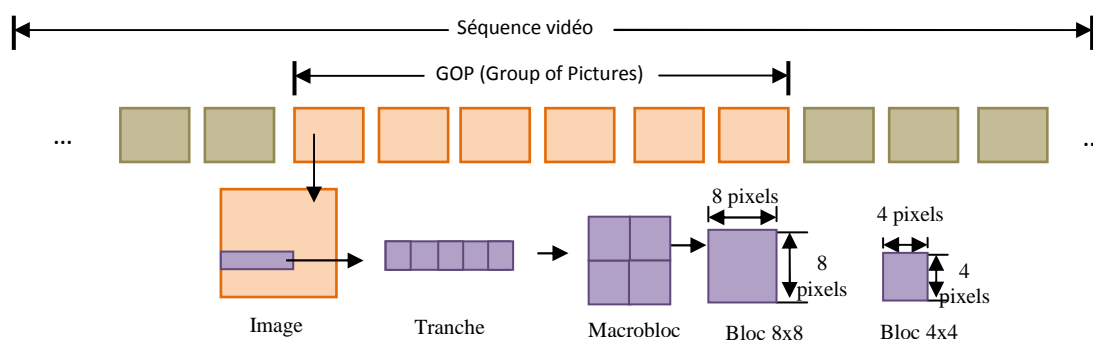


Figure 1.8 : Hiérarchie des données dans le flux vidéo.

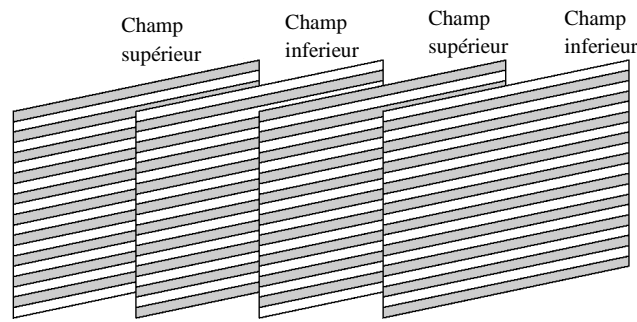


Figure 1.9 : Séquence vidéo entrelacée [2].

1.3.3 La prédiction

1.3.3.1 La prédiction Intra

Pour profiter de la propriété de corrélation spatiale dans une image, le standard H.264 fait appel à une prédiction Intra sur des blocs 16x16 ou 4x4. En effet, le macrobloc d'intérêt peut être prédit à partir des macroblocs généralement situés en dessus et à gauche étant donné qu'ils ont été déjà prédits. La différence entre le macrobloc réel et sa prédiction est ensuite codée offrant ainsi une représentation avec un nombre de bits réduit par rapport à celui utilisé pour la transformation directe du macrobloc lui-même.

Il existe deux modes de codage Intra :

- le mode Intra 4x4 : neuf modes de prédiction sont possibles pour coder un bloc 4x4 de luminance. La prédiction est faite par extrapolation ou interpolation des échantillons situés en haut ou à gauche selon la direction du mode (voir chapitre 4) ;
- le mode Intra 16x16 qui est recommandé dans le cas des zones d'image régulières. Dans ce cas quatre modes de prédiction existent pour les échantillons de luminance, tel qu'illustré à la figure 1.10.

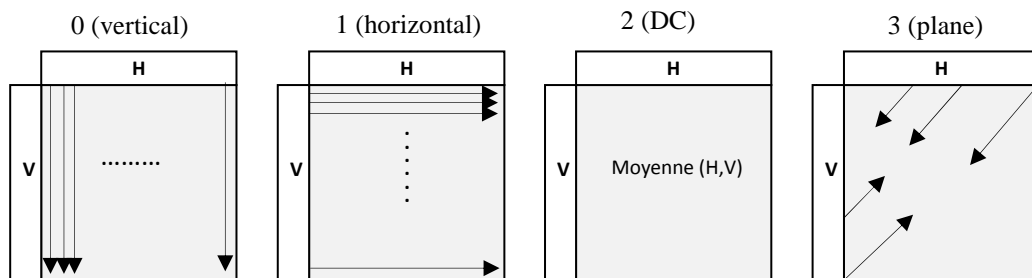


Figure 1.10 : Les modes de prédiction intra 16x16.

La prédiction des échantillons de chrominance d'un macrobloc repose sur une technique semblable à celle utilisée pour la luminance des macroblocs Intra 16x16. La prédiction intra au-delà des limites de la tranche n'est pas autorisée afin de maintenir l'indépendance des tranches les unes par rapport aux autres.

- La prédiction basée sur des blocs de taille 8 x 8 pixels, qui est essentiellement identique au mode de prédiction 4x4 mais utilisé uniquement dans le profil H.264/AVC haute définition (voir section 1.3.7).

1.3.3.2 La prédiction Inter

La prédiction Inter est basée sur les techniques d'estimation et de compensation de mouvement. Ces techniques permettent de tirer profit des redondances temporelles qui existent entre les images successives. Leur principe de fonctionnement consiste à mesurer d'abord le déplacement d'objets entre images successives (estimation du mouvement). La compensation de mouvement est basée sur la création de l'image prédite et la déduction de l'image résiduelle (l'erreur) qui sera transmise avec la représentation des mouvements à la place de l'image elle-même.

La norme H.264/AVC introduit la notion de la compensation en structure d'arbre (*Tree-structured compensation*), c'est à dire que si le macrobloc est partitionné en blocs de 8x8, chacun de ces blocs peut être divisé en sous-blocs de taille inférieure (8x4, 4x8 ou 4x4) tel qu'illustré à la figure 1.11. Le vecteur de mouvement et le choix de la partition doivent être codés et transmis dans le flux de la vidéo compressée. L'intérêt de cette structure est d'augmenter la performance de la compression en associant les partitions de grandes tailles aux zones homogènes et les plus petites aux zones détaillées.

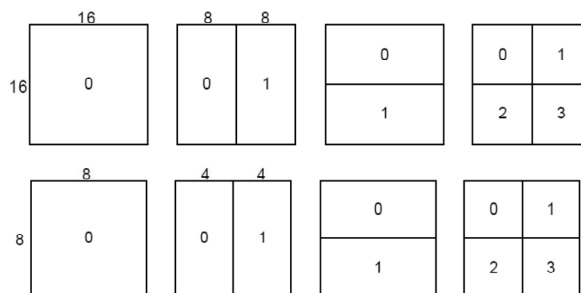


Figure 1.11 : Partitions (16x16, 8x16, 16x8, 8x8) et sous partitions (8x8, 4x8, 8x4, 4x4) du macrobloc.

Chaque partition ou sous-partition du macrobloc codé en intra est prédite à partir d'une zone de la même taille dans une image de référence. Le vecteur de mouvement peut être calculé à une précision du quart de pixel pour la luminance et du huitième du pixel pour la compensation de chrominance. Dans ce cas, les résolutions sont obtenues par interpolation sur les pixels existants, ce qui permet l'amélioration de la précision de calcul du mouvement, et par conséquent une réduction du débit.

La figure 1.12 représente des exemples de vecteurs de mouvement. Contrairement au cas de la figure 1.12(b) dans lequel la prédiction renvoie à des échantillons existants, le cas de la

figure 1.12(c) nécessite une génération des échantillons par interpolation entre des échantillons adjacents dans l'image de référence.

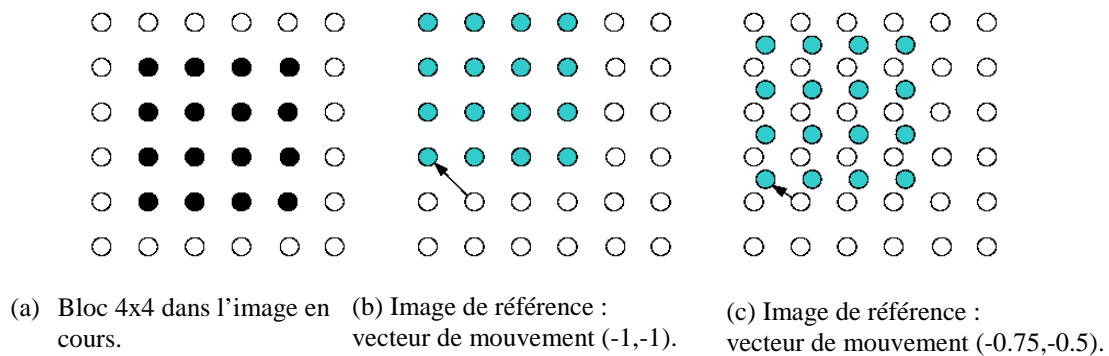


Figure 1.12 : Exemples de vecteurs de mouvements.

Les vecteurs de mouvement font également l'objet d'un codage différentiel, basé sur une prédiction à partir des vecteurs des blocs voisins. Leur prédiction ne peut jamais intervenir au-delà des limites de la tranche.

1.3.4 Transformation et quantification

Cette étape permet de réduire la redondance spatiale de l'erreur de prédiction. La norme H.264/AVC utilise la transformée DCT entière pour tous les blocs de données résiduelles, en utilisant la matrice H1 de taille 4x4. Dans le cas de l'application du mode Intra 16x16, la transformée de Hadamard est en plus utilisée pour des matrices 4x4 des coefficients DC (*Direct Current*) de luminance, en utilisant la matrice H2. Une transformée de Hadamard est également utilisée pour transformer les quatre coefficients DC de chaque composante de chrominance. Dans ce cas, la matrice H3 de taille 2x2 est utilisée. Les matrices H1, H2 et H3 sont indiquées comme suit [3] :

$$H1 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix}, \quad H2 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix}, \quad H3 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

L'utilisation de la DCT entière permet de réduire la complexité de calcul en utilisant seulement des opérations de décalage, d'addition et de soustraction. Dans le cas de la transformée de Hadamard, on utilise seulement les opérations d'addition et de soustraction. De plus, la transformée inverse est réalisée par des opérations sur des entiers exactes, les discordances de transformée inverse, causées dans les standards précédents, sont ainsi évitées.

Pour quantifier les coefficients de la transformée, la norme H.264/AVC utilise 52 pas de quantification sélectionnés pour chaque macrobloc par le paramètre de quantification (QP) défini dans une progression logarithmique, ce qui permet une plus grande finesse.

Les coefficients quantifiés d'un bloc sont en général analysés en zigzag et transmis par codage entropique.

La figure 1.13 illustre l'ordre de transmission des données dans un macrobloc. Pour le mode de transmission Intra 16x16, le bloc « -1 » est transmis en premier, il contient les coefficients DC transformés dans chaque bloc luma 4x4. Les blocs résiduels de 0 à 15, qui contiennent les coefficients AC (*Alternative Current*), sont transmis dans l'ordre indiqué. Les blocs 16 et 17 contenant les coefficients DC transformés dans chaque bloc chroma, et enfin, les blocs 18 à 25 sont transmis (sans les coefficients DC déjà transmis).

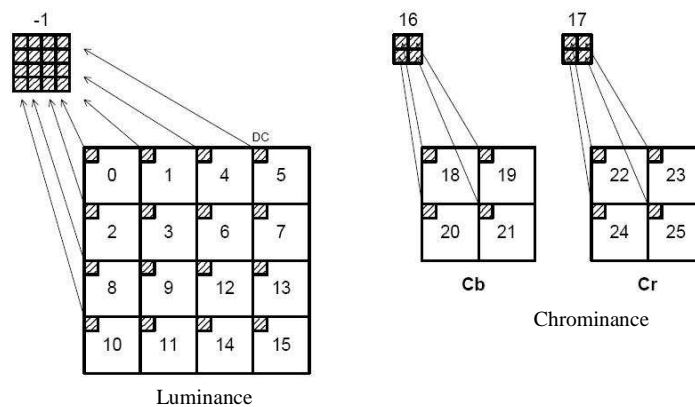


Figure 1.13 : Ordre de transmission des blocs résiduels dans le mode Intra 16x16.

1.3.5 Filtre anti-bloc adaptatif

Le codage vidéo basé sur le découpage en blocs produit des artefacts. L'application du filtre anti-bloc aux images servant de référence, permet de réduire la pixellisation et améliorer la qualité de la vidéo. Toutefois, il est important de distinguer entre les bords réels de l'image et ceux générés par certaines étapes du codage (prédiction, transformation et quantification). Selon la nature du bord, le filtre est appliqué en utilisant un paramètre Bs (*Boundary-strength*) ayant une valeur entière entre 0 et 4 qui indiquent la force du filtrage.

Deux seuils α et β sont définis pour détecter des contours réels et décider d'appliquer le processus de filtrage. α et β dépendent essentiellement de la valeur moyenne du paramètre de quantification QP des deux blocs 4x4 voisins séparant le bord à filtrer.

Un groupe d'échantillons ($p_2, p_1, p_0, q_0, q_1, q_2$), représenté à la figure 1.14, est filtré si :

$$Bs > 0 \text{ et } |p_0 - q_0| < \alpha \text{ et } |p_1 - p_0| < \beta \text{ et } |q_1 - q_0| \leq \beta.$$

Le processus de filtrage est décrit avec plus de détails dans [3].

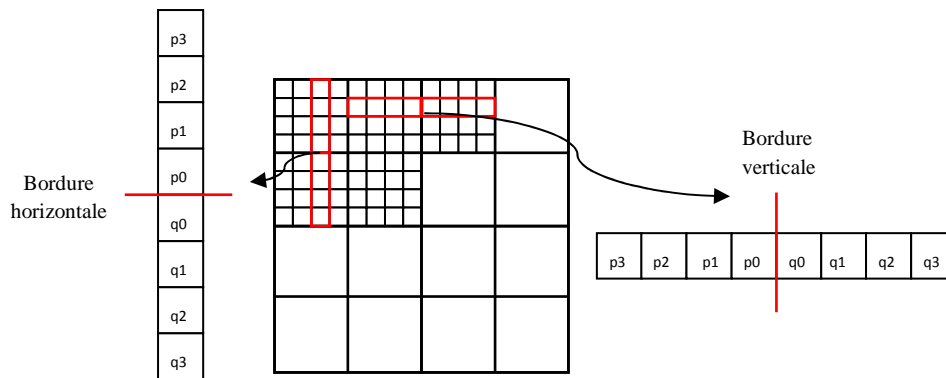


Figure 1.14: Illustration des bordures horizontales et verticales.

1.3.6 Codage entropique

Le codage vidéo produit un nombre de valeurs qui doivent être codées pour former le flux compressé. Parmi ces valeurs :

- les coefficients transformés quantifiés ;
- des informations permettant la création de la prédiction ;
- des informations sur la structure des données compressées et les outils de compression utilisée dans l'encodage ;
- des informations sur la séquence vidéo complète.

La norme H.264/AVC utilise deux méthodes de codage entropique : Le codage VLC (*Variable Length Coding*) et le CABAC (*Context-Adaptive Binary Arithmetic Coding*).

1.3.6.1 Le codage VLC

Le codage entropique à longueur variable (VLC) associe aux événements un mot de code. Les codes ont des longueurs variables et un nombre entier de bits. Nous présentons dans cette sous-section les codages VLC utilisés par le le H.264/AVC, à savoir le Golomb exponentiel et le CAVLC (*Context-Adaptive Variable Length Coding*).

a) Le codage Golomb exponentiel

A l'exception des coefficients transformés quantifiés qui utilisent un codage CAVLC, les autres unités utilisent un codage Golomb exponentiel (ou Exp-Golomb). Il s'agit d'un ensemble unique et illimité de mots-codes défini pour tous les éléments syntaxiques.

Le tableau 1.1 représente quelques mots-codes. La structure de chaque mot est la suivante :

[M zéros] [1] [INFO]

Tableau 1.1. Mots –codes Golomb exponentiel.

Valeur	Mot- code
0	1
1	010
2	011
3	00100
4	00101
5	00110
6	00111
7	0001000
8	0001001
...	...

INFO est constitué de M bits porteurs d'information. La longueur de chaque mot code est $(2M+1)$ bits. Pour le mot code 0001001, par exemple, $M=3$ et $Info = 1$. Lors du décodage, après la lecture de M et INFO, la valeur à décoder est déduite par : $Valeur = 2^M + INFO - 1$.

b) Le codage CAVLC

La méthode CAVLC applique un schéma plus sophistiqué pour transmettre les coefficients de la transformée quantifiés. Ce codage a été conçu pour exploiter les caractéristiques des blocs 4x4 quantifiés. Le choix des tables de codage utilisées dépend des coefficients non nuls des blocs voisins précédemment codés.

Prenons à titre d'exemple le bloc 4x4 représenté à la figure 1.15.

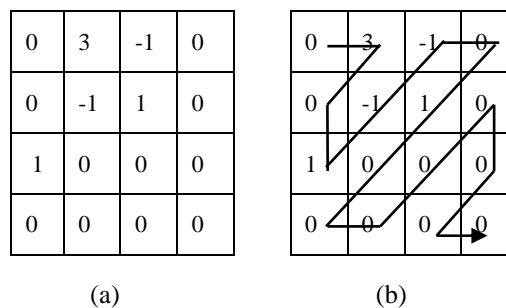


Figure 1.15 : (a) Exemple de bloc 4x4, (b) Lecture en zigzag des coefficients.

Après un balayage en zigzag, nous obtenons : 0, 3, 0, 1, -1, -1, 0, 1, 0,...

Le codage CAVLC suit les étapes suivantes :

- Codage du nombre de tous les coefficients non nuls (CoeffTotal) et les *trailing ones* (T1) : il s'agit de coder le nombre total des coefficients non nuls, variant de 0 à 16 dans un bloc 4x4, ainsi que le nombre des T1. Les T1 sont les derniers ± 1 , leur nombre ne dépasse pas 3. Dans notre exemple, le nombre de CoeffTotal est égal à 5 et le nombre de T1 est égal à 3 (le premier chiffre 1 est considéré comme un coefficient non nul);
- Codage des signes des T1 avec un seul bit;
- Codage des niveaux des coefficients non nuls restant en commençant par les hautes fréquences (dans notre exemple +1 et +3) ;

- Codage du nombre total N_0 de zéro 0 se positionnant avant le dernier coefficient non nul ($N_0=3$);
- Codage des longueurs de séquences de zéro se trouvant avant chaque coefficient non nul en commençant par les plus hautes fréquences (Par exemple, pour les cinq coefficient non nuls de la figure 1.15 (b) nous avons : run-before(5)=1, run-before(4)=0, run before(3)=0, run before(2)=1 , run before(1)=1).

1.3.6.2 Le codage CABAC

L'efficacité du codage entropique peut encore être améliorée avec le codage adaptatif arithmétique CABAC. Le codage CABAC consiste essentiellement en trois étapes [14] :

- la binarisation ;
- le « Context Modeling » ;
- le codage binaire arithmétique.

L'étape de binarisation consiste à convertir les éléments de syntaxe⁶ non binaire en chaîne binaire appelée « bins ». Les éléments de syntaxe non binaire sont les éléments pouvant avoir plus de 2 valeurs, tels les vecteurs de mouvement et les coefficients de la transformation. La modélisation de contexte consiste à choisir un modèle pour chaque « bin » avant d'appliquer le codage arithmétique en fournissant des estimations des probabilités conditionnelles relatives aux symboles en se basant sur les symboles voisins codés précédemment.

Le codage arithmétique est basé sur la subdivision récursive d'intervalle. En appliquant le codage arithmétique, chaque bin est codé séquentiellement selon le modèle de contexte associé pour produire un nombre appelé offset. Le modèle de contexte n'est pas fixe, mais renouvelé pour chaque encodage de bin.

Ce codeur est beaucoup plus lent que le précédent, mais il permet de gagner de 10 à 15 % par rapport au CAVLC.

1.3.7 Les profils et niveaux

Le profil définit un ensemble d'outils de codage ou d'algorithmes qui peuvent être utilisés pour générer un flux compatible, tandis qu'un niveau impose des contraintes à certains paramètres clés du flux.

Les profils correspondent à une approche qualitative : ils regroupent les outils de codage et les algorithmes les plus appropriés pour diverses situations d'utilisation : diffusion, production...

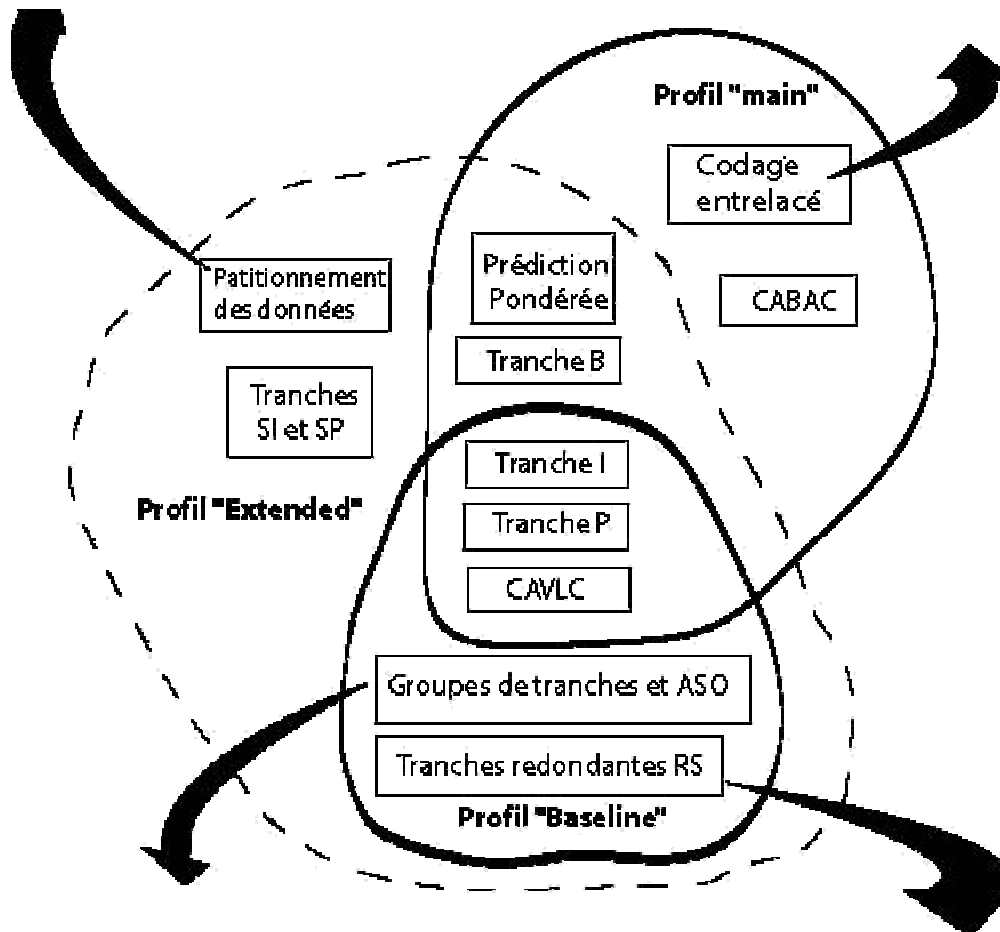
⁶ Un élément de syntaxe est un élément de données dans le flux vidéo.

Les profils définissent ainsi les types d'images qui seront utilisés (Images I, P, B...), les procédés de codage employés, etc. Les décodeurs n'auront pas à implémenter tous les outils de la norme mais seulement ceux destinés à l'usage visé.

La norme H.264/AVC prend en charge les profils suivants : Le profil de base (*baseline*), Le profil principal (*main*) et le profil étendu X (*extended*). La figure 1.16 représente les fonctions supportées par chaque profil.

Le partitionnement des données permet de séparer les éléments de syntaxe d'importance plus ou moins élevée dans des paquets de données pour leur offrir le niveau de protection adéquat et améliorer la fiabilité du flux.

Codage en Picture-adaptive frame-field (PAFF): le codage de l'image s'effectue en une seule trame ou par entrelacement (deux champs). De même pour le codage en Macroblock-adaptive frame-field (MBAFF), le codage de deux macrobloccs appariés se fait en macrobloccs 16x16 ou en entrelaçant les champs des deux macrobloccs.



Parmi les fonctions qui permettent d'améliorer la résistance aux erreurs :

- La fonction d'ordonnancement arbitraire des tranches (ASO) qui autorise un décodage arbitraire des tranches.
- Les groupes de tranches qui peuvent être décodés indépendamment. Ils sont formés d'un ordonnancement flexible de macrobloccs (FMO). Ces tranches peuvent être horizontales, verticales, rectangulaires, entrelacées ou dispersées sur toute l'image.

Transmission d'une version additionnelle de tout ou d'une partie de l'image dans une qualité moindre pour améliorer la résistance aux erreurs et aux pertes.

Figure 1.16 : Profils de la norme H.264/AVC.

Chaque profil comporte différents niveaux. Ils apportent une approche quantitative, décrivant différents paliers de performance pour un profil donné et correspondant à autant de degrés de complexité dans la mise en œuvre.

Le profil de base a été conçu pour réduire au minimum la complexité et fournir plus de fiabilité et de souplesse d'utilisation sur un large éventail de conditions et d'environnements de réseau. Quant au profil principal, il améliore l'efficacité du codage. Le profil étendu a été conçu pour combiner entre les deux profils précédents afin d'assurer fiabilité et efficacité de codage et ajouter des modes utiles aux applications telles que le streaming vidéo [15].

Le projet JVT a connu une extension visant à répondre aux besoins des applications les plus exigeantes en termes de résolution de vidéo. Cette extension, connue sous le nom de FRExt (Fidelity Range Extensions), a été finalisée en Juillet 2004 et a produit quatre nouveaux profils appelés collectivement profil haut ou de haute définition (*High profiles*) [15]. En 2007, une autre extension a été approuvée "le codage vidéo évolutif", qui contient trois profils additionnels: le profil évolutif de base, le profil évolutif haut, et le profil évolutif Intra haut. L'objectif du codage vidéo évolutif est de permettre une reconstruction efficace de la vidéo à différents débits et taux d'images et résolution d'affichage du même flux [16].

1.4 Conclusion

Dans ce chapitre nous avons présenté brièvement les différents modules qui composent le codec H.264/AVC, ainsi que les outils qui ont permis de renforcer l'efficacité du codage par rapport aux codecs antérieurs.

En effet, le domaine de la compression vidéo connaît un intérêt croissant avec le développement des réseaux et des applications multimédias. Toutefois, la disponibilité de la vidéo à travers Internet et des œuvres multimédias particulièrement a motivé la recherche de solutions qui permettraient essentiellement de protéger les œuvres numériques. Le tatouage fait partie de ces outils qui peuvent être intégrés dans les systèmes de gestion de droits numériques (DRM) pour être exploités par les fournisseurs de contenu.

Dans le prochain chapitre nous présentons des généralités sur les techniques de tatouage, leur application à la vidéo et particulièrement au codec H.264/AVC.

2.1 Introduction

Les techniques de tatouage ont été développées pour proposer de nouvelles méthodes de protection et de gestion des médias numériques. L'intérêt pour ces techniques est essentiellement dû au développement des technologies de l'information et de la communication qui a facilité la distribution des documents multimédias numériques.

La vidéo attire particulièrement la communauté des chercheurs dans ce domaine vu le succès qu'elle connaît actuellement et les enjeux qu'elle porte avec la prolifération des serveurs de vidéos.

Le tatouage vidéo décrit les méthodes d'insertion des informations supplémentaires dans la vidéo numérique dans le but de répondre à des applications telles que la protection des droits d'auteurs, l'authentification et les communications secrètes.

Aujourd'hui, les activités de recherche dans ce domaine portent sur la norme vidéo H.264/AVC. Les chercheurs tentent d'adapter les méthodes existantes ou de concevoir de nouvelles méthodes spécifiques à ce codec. Toutefois, sur le plan technique, le taux de compression élevé qui est offert par cette norme constitue un défi pour l'insertion de la marque.

Dans ce chapitre nous présentons des généralités sur les techniques de tatouage, leur application à la vidéo, et un état de l'art sur les techniques de tatouage appliquées au codec H.264/AVC.

2.2 Généralités sur les techniques de tatouage vidéo

2.2.1 Définition du tatouage

Le tatouage (ou le marquage) numérique est l'art de dissimuler l'information d'une manière robuste et invisible [17]. Les techniques de tatouage s'apparentent à la stéganographie. Cette dernière consiste à dissimuler un message dans un autre, essentiellement pour tenir une communication secrète. La lecture du message caché nécessite la connaissance de la procédure de dissimulation.

Les techniques de tatouage ont été proposées initialement pour protéger les documents multimédias en insérant des données (appelées marque ou signature), qui identifieraient le propriétaire du document et seraient une preuve de la paternité de l'œuvre. Contrairement à la cryptographie (ou particulièrement au chiffrement), le tatouage n'empêche pas un utilisateur de voir (ou d'écouter) un document multimédia, et les données insérées dans le document hôte ne doivent pas l'altérer. En plus de la protection des droits d'auteurs, plusieurs applications du tatouage sont apparues par la suite. Des exemples sont cités dans la sous-section suivante.

2.2.2 Applications et critères

Les techniques de tatouage peuvent être exploitées dans plusieurs applications telles que [17] :

- la protection des droits d'auteurs (contrôler la paternité des documents) ;
- les contrôles de copies (empêcher les copies non autorisées) ;
- la surveillance de diffusion (identifier et contrôler les vidéos diffusées) ;
- l'authentification de la vidéo (s'assurer de l'origine du document (authenticité) et que le contenu original n'a pas été modifié (intégrité)) ;
- l'amélioration du codage vidéo, en insérant des informations supplémentaires pour la correction d'erreurs ;
- la publicité (vérifier la fréquence d'affichage d'une publicité) ;
- l'archivage (ajouter des métadonnées telles que le nom du propriétaire, la date, etc. pour l'archivage) ;
- l'empreinte digitale (détecter les utilisateurs malveillants qui diffusent illégalement un document en y insérant par exemple des numéros qui identifient les utilisateurs).

Chaque application a ses propres exigences par rapport à certains critères tels que la capacité du marquage, la robustesse et l'imperceptibilité (aussi appelée fidélité).

"La capacité de marquage" est la quantité d'information que l'on insère dans le document hôte. Certaines applications telles que les communications secrètes nécessitent des capacités de marquage relativement importantes.

"La sécurité du tatouage" est liée à la confidentialité des données insérées. Ce critère est souvent confondu avec "la robustesse" qui permet de rendre possible de déceler la marque malgré les différentes attaques possibles, à moins que le document hôte devienne inexploitable. En effet, la robustesse est indispensable pour les applications de protection des droits d'auteurs, toutefois, dans certains cas, "la fragilité" de la marque est préférable s'il s'agit de faire, à titre d'exemple, un contrôle d'intégrité. Dans ce cas, la marque sera altérée ou supprimée à la moindre modification du document hôte.

Quant à "la réversibilité du marquage", c'est une caractéristique qui offre, aux personnes autorisées, la possibilité de récupérer le document original après marquage.

La difficulté dans la conception d'un système de tatouage est de réaliser un compromis entre les différents critères, souvent difficiles à satisfaire en même temps, notamment lorsqu'il s'agit de la robustesse, de l'imperceptibilité et de la capacité d'insertion.

2.2.3 Quelques techniques de tatouage

En 1990, Tanaka et al.[18] ont proposé l'idée d'insérer des données secrètes dans une image, dans le bruit de dithering⁷. En 1993, Tirkel et al. [19] ont proposé l'utilisation de la méthode des bits les moins significatifs (LSBs) qui consiste à remplacer les bits les moins significatifs des pixels par les bits du message secret. Deux années plus tard, Bender et al.[20] ont proposé la méthode du Patchwork qui permet de cacher un bit seulement dans une image de manière robuste. L'insertion se base sur le choix aléatoire de paires de pixels, dont les luminances des premières valeurs des paires sont incrémentées de 1 et les deuxièmes sont décrémentées de 1. La méthode de l'étalement spectral est apparue dans la même année (1995) pour apporter de la robustesse aux attaques [21][22]. En effet, la marque à insérer est représentée par une séquence binaire pseudo aléatoire qui sera ajoutée par bloc à l'image, tel qu'illustré à la figure 2.1, entraînant la modification des LSBs. La détection de la marque se fait par démodulation.

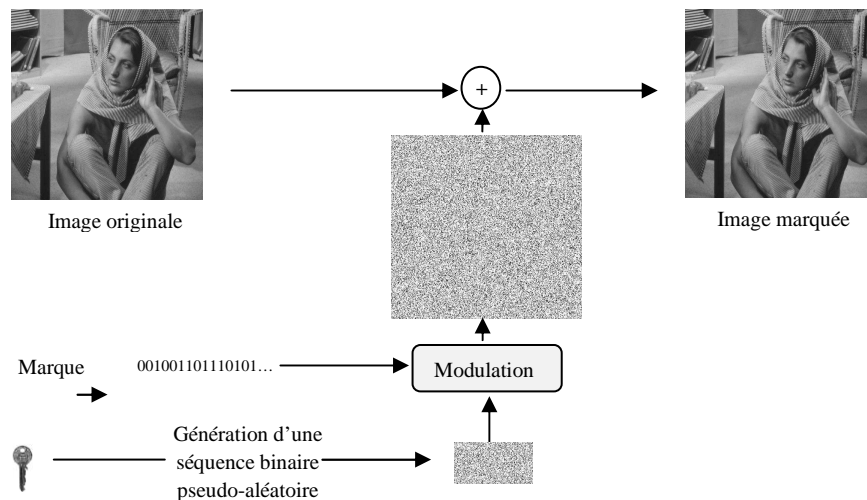


Figure 2.1 : Insertion de la marque basée sur la méthode de l'étalement du spectre.

Plus tard, d'autres schémas sont apparus pour apporter des améliorations à la robustesse et à la qualité visuelle en exploitant les propriétés du système visuel humain, comme l'exploitation de la composante de chrominance bleue pour laquelle l'œil est moins sensible [23].

Les méthodes de tatouage par quantification quantifient le signal hôte selon un dictionnaire prédéfini de quantificateurs, correspondant à différents messages. Le décodage détermine la grille de quantification la plus proche de la composante à marquer. Parmi les techniques quantificatives les plus connues : La Modulation d'Indices de Quantification (QIM) [24].

Dans la méthode QIM, le signal hôte est modifié de façon à correspondre à un état de quantification (un mot de code), selon le message à insérer. Le quantificateur est constitué d'un

⁷ Le dithering consiste à rajouter un bruit à un signal pour empêcher l'apparition d'artéfacts dus à la quantification.

ensemble d'états de quantification possibles, il représente un dictionnaire structuré, réparti uniformément dans l'espace. Le nombre de symboles à insérer correspond au nombre de quantificateurs. Dans le cas d'un message binaire deux états de quantification sont nécessaires.

2.2.4 Domaine d'insertion

Parmi les méthodes de tatouage d'images, on distingue celles qui opèrent dans le domaine spatial et qui sont généralement simples à implémenter, et celles qui opèrent dans le domaine des transformées (DCT, DWT, etc.), et qui nécessitent la transformation du signal hôte par la transformée de Fourier (FFT), la transformée en Cosinus, la transformée en Ondelettes ou autre. Ces méthodes sont plus complexes que les premières mais assurent une meilleure robustesse et imperceptibilité du marquage [25].

2.2.5 Les techniques de tatouage appliquées à la vidéo

Concernant les techniques de tatouage vidéo, elles suivent en général les techniques utilisées pour l'image, étant donné que la vidéo est considérée comme une succession d'images. Toutefois, il est important de prendre en compte les propriétés de la séquence vidéo pour l'amélioration des performances. En effet, la grande taille de la vidéo et particulièrement son axe temporel peut être exploité pour intégrer la marque, même si certaines contraintes doivent être prises en compte comme indiqué dans la section 2.3.

La figure 2.2 représente une classification possible des techniques de tatouage vidéo en fonction du domaine d'insertion.

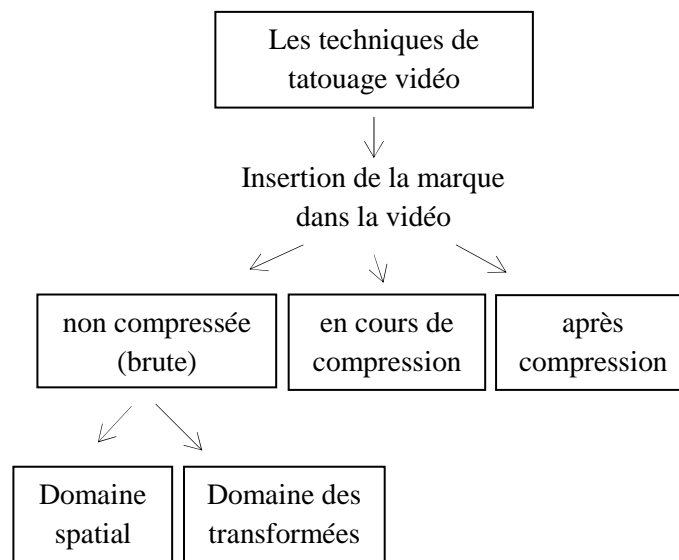


Figure 2.2 : Classification des méthodes de tatouage vidéo.

Dans le domaine fréquentiel, les schémas proposés exploitent les moyennes et basses fréquences pour insérer une marque robuste.

Un exemple de méthode de tatouage dans le domaine non compressé est présenté par Park *et al.* [26]. Il est basé sur l'utilisation des coefficients AC du domaine 3D-DCT et les caractéristiques vidéo telles que la quantité d'informations de texture et de mouvement. Cette méthode est robuste à des attaques comme la compression MPEG, insertion image, perte ou permutation d'images.

L'utilisation de la vidéo non compressée pour le tatouage nous offre le choix d'effectuer des compressions selon différents débits. Cependant, la marque doit être robuste aux divers formats de compression, et puisqu'il n'est pas toujours possible et pratique de transmettre et stocker des vidéos numériques brutes, la plupart des travaux des chercheurs sont effectués dans le domaine compressé comme présenté dans les sous-sections suivantes où des exemples de tatouage du codec vidéo H264/AVC sont présentés.

2.2.6 Les attaques

L'attaque est considérée comme étant tout traitement pouvant altérer la marque et causer une ambiguïté dans son extraction. Il existe différents types d'attaques classifiées par Doërr *et al.* [27] selon qu'elles traitent la robustesse ou la sécurité, tel qu'illustré à la figure 2.3.

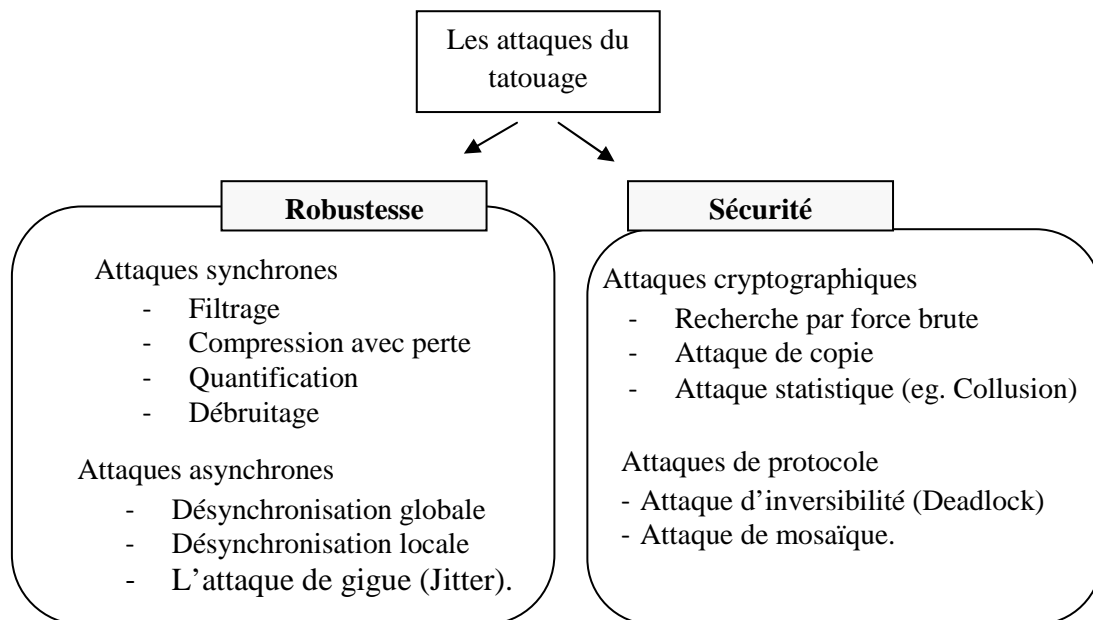


Figure 2.3 : Les attaques du tatouage.

Dans le contexte du tatouage vidéo, la collusion est considérée comme l'une des questions difficiles, souvent utilisée pour évaluer la sécurité du tatouage. Elle a été mentionnée en premier lieu dans le contexte des applications à empreintes (*fingerprint*) utilisées pour détecter l'origine d'une copie illégale. L'attaque consiste à combiner des informations de différentes copies afin d'éliminer ou d'atténuer la marque intégrée.

Pour le cas de la vidéo, chaque image est considérée comme un contenu pouvant être marqué individuellement. Pour fournir un marquage continu et faire face à certaines attaques telles que la suppression d'images, il est recommandé de marquer toutes les images de la vidéo. De ce fait, la vidéo peut subir une attaque de collusion si les différentes images sont collectées et combinées pour supprimer la marque [28].

Dans le cas de la vidéo compressée, l'une des attaques les plus connues consiste à effectuer une deuxième compression (ou plus) de la vidéo pour rendre la marque indétectable. Toutefois, plusieurs ré-encodages de la vidéo pourraient altérer sa qualité visuelle et la rendre inexploitable.

2.3 Les techniques de tatouage appliquées au standard H.264/AVC

2.3.1 Les exigences du tatouage vidéo

Pour la vidéo compressée, il s'avère plus difficile d'insérer la marque dans le flux binaire en raison des contraintes supplémentaires nécessaires. En effet, pour fournir un tatouage continu, il est recommandé d'insérer la marque dans toutes les images de type I, P et B tout en veillant à satisfaire certaines recommandations énumérées ci-dessous [28]:

- l'insertion de la marque doit préserver le débit afin de ne pas aller à l'encontre du principe de la compression ;
- L'imperceptibilité de la marque doit répondre aux exigences de l'application et du type d'images (photographiques, médicales, cartographiques, etc.) ;
- Le taux d'insertion fourni par la méthode doit être suffisant pour contenir les informations nécessaires ;
- La méthode de tatouage doit assurer un décodage correct de la vidéo et une extraction fiable de la marque ;
- L'augmentation du temps de calcul introduite par le programme d'insertion de la marque doit rester négligeable.
- La méthode de tatouage doit assurer la sécurité des données insérées ainsi que la robustesse contre les éventuelles attaques ;

2.3.2 La zone d'insertion

Les possibilités de marquage du codec H.264/AVC semblent être très limitées étant donné le taux de compression élevé introduit par les nouvelles techniques de codage. Les schémas proposés sont souvent basés sur les vecteurs de mouvement ou bien sur les coefficients de la DCT quantifiés ou bien la combinaison entre les deux zones d'insertion pour fournir une solution complète de tatouage fragile et robuste. Les techniques apparues plus récemment exploitent les modes de prédiction. La différence entre les méthodes réside dans le choix des zones d'insertion qui s'adaptent le mieux aux propriétés du système visuel humain. L'objectif serait de

minimiser la visibilité de la marque tout en satisfaisant également les autres contraintes (débit, capacité d'insertion, etc.).

2.3.2.1 Insertion de la marque au cours du processus de compression

a) Insertion de la marque dans les vecteurs de mouvement

Les vecteurs de mouvement ont été proposés pour porter la marque en 1997 [29]. En se basant sur le principe de parité, une légère modification est apportée aux vecteurs de mouvement. D'autres méthodes ont proposé d'exploiter les amplitudes relativement élevées des vecteurs car elles concernent les mouvements rapides où les détails sont imperceptibles [30].

Dans le cas du standard vidéo H.264/AVC, l'estimation de bloc de mouvement de tailles variables pourrait fournir plus de capacité pour intégrer la marque. La plupart des travaux présentés pour ce codec et qui utilisent les vecteurs de mouvement concernent l'insertion d'une marque fragile qui pourrait être utilisée pour la détection des zones altérées dans une vidéo.

Kuo et al.[31] ont proposé de restreindre la zone de recherche du meilleur appariement durant le processus de marquage. Par exemple, si nous considérons un intervalle de recherche divisé en deux régions I et II, et le macrobloc se dirige vers la région I, la direction du mouvement est modifiée vers II (mode inter) pour insérer le bit "0", toutefois le mode de codage est changé en intra s'il n'y a pas un bon appariement de macrobloc dans cette zone ; la fonction *Rate-distortion-cost* (voir chapitre 4, section 4.4) est utilisée pour sélectionner le meilleur mode de codage qui permettrait de maintenir la qualité de la vidéo.

b) Insertion de la marque dans les coefficients DCT

En plus des vecteurs de mouvement, les coefficients de luminance AC quantifiés peuvent également être utilisés pour un marquage fragile ou semi⁸ fragile. Chen et al. [32] ont proposé un schéma de tatouage semi fragile qui permet de détecter la zone illégalement modifiée de la vidéo. Cette application peut être utilisée dans les systèmes de vidéo-surveillance. Dans le même contexte, de nouvelles idées proposent de profiter des avantages du codage entropique et en particulier du CAVLC afin de maintenir le débit tout en intégrant la marque. Cela peut être fait, par exemple, en veillant à ce que le choix de la table de correspondance, utilisée pour encoder les coefficients non nuls et les *trailing-ones*, reste inchangé [33].

Les coefficients de luminance quantifiés sont également utilisés pour insérer une marque robuste après l'étape de quantification qui est une opération avec perte. En effet, en considérant un bloc issu d'une erreur de prédiction des images I, P ou B, ce bloc est transformé et quantifié pour

⁸ Le tatouage semi-fragile est un tatouage qui résiste à des manipulations autorisées.

obtenir les coefficients qui sont utilisés pour insérer une marque. L'étape d'extraction est effectuée après l'opération de décodage entropique du décodeur.

Les coefficients sont sélectionnés aléatoirement ou expérimentalement parmi les coefficients AC et en prenant en considération la direction du balayage, tel qu'illustré à la figure 2.4. En effet, les informations insérées au niveau des hautes fréquences pourraient causer une augmentation du débit et les informations insérées en basses fréquences pourraient conduire à une dégradation de la vidéo [34].

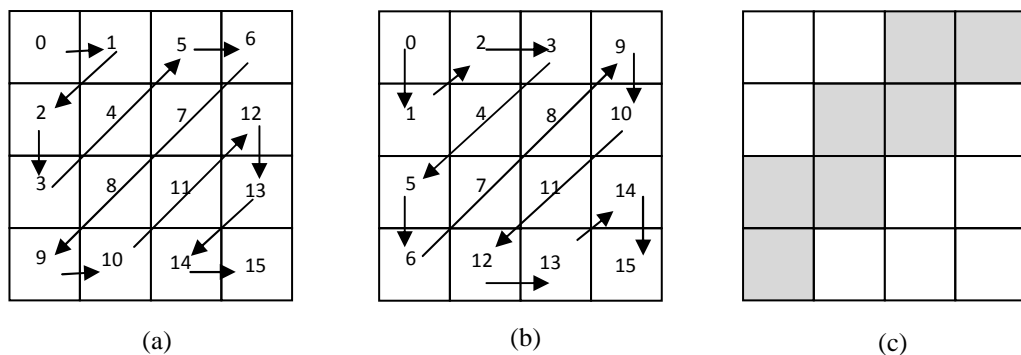


Figure 2.4 : (a) Le balayage en Zig-zag, (b) Le balayage en champs, (c) Exemple de zones d'insertion dans un bloc de coefficients DCT.

En pratique, l'insertion se fait généralement dans les images de type I, car P et B sont fortement compressées par les techniques d'estimation et de compensation de mouvement.

L'utilisation des macroblocs intra-prédit 16x16 est évitée, car cette prédiction concerne les régions homogènes et l'insertion de la marque causerait des artefacts visibles à ce niveau [4], bien que Golikeri et al. [5] ont suggéré l'utilisation de coefficients de Hadamard (mode 16 x 16) pour intégrer un tatouage robuste, car ils contiennent la plus grande énergie et assureraient donc la robustesse du marquage.

La sélection des blocs serait basée de préférence sur des modèles qui analysent le contenu de la vidéo et exploitent les fonctions vidéo pour une meilleure sélection des zones d'intégration. Chen et al. [35] ont proposé une méthode pour insérer la marque dans les blocs à l'aide de deux algorithmes pour des blocs de haute et basse énergie. L'entropie est aussi utilisée dans [36] pour la sélection des blocs.

Seo et al. [37] ont utilisé dans leur schéma le processus CABAC et les LSBs afin de protéger la propriété du contenu de la vidéo. L'idée est de sélectionner, d'une part, le bloc pour lequel il est très possible d'avoir de nombreux changements, car ils se réfèrent aux régions de hautes fréquences caractérisées par une imperceptibilité élevée. D'autre part, dans chaque macrobloc sélectionné, le coefficient de la plus grande valeur des basses fréquences (à l'exception du DC)

est choisi pour intégrer un bit de la marque. Ce choix est déterminé par les contextes définis dans le processus de CABAC. La technique présentée est robuste contre les attaques de type modifications géométriques, recadrage, flou, etc.

c) Insertion de la marque en exploitant les IPCM et les modes de prédiction

Kapotas et al.[38] ont introduit l'utilisation des macroblocs codés en IPCM (Intra pulse code modulation). Le codage des macroblocs en IPCM permet de garder la valeur initiale des échantillons et ne pas subir une prédiction inter ou intra, l'objectif est de préserver la qualité visuelle initiale de certaines régions. Dans cette méthode, pour augmenter la capacité d'insertion il est nécessaire de forcer des macroblocs au codage IPCM, ce qui cause une augmentation importante du débit.

Par ailleurs, quelques méthodes de tatouage basées sur les modes de prédiction ont été proposées. La méthode proposée par Hu et al. [7] consiste à modifier les modes de prédiction intra 4x4 en utilisant une table de correspondance. Afin d'améliorer la capacité d'insertion et de contrôler l'augmentation du débit, les règles de mise en correspondance ont été améliorées par Yang et al. [6]. Dans ce schéma, chacun des trois blocs de 4x4 intra qui répondent aux exigences d'insertion de la marque constituent un groupe. Chaque deux bits successif de la marque sont modulés aux modes de ces trois blocs (voir chapitre 4, tableau 4.1) et seulement un bloc change sa valeur de prédiction pour l'insertion de ces deux bits. Dans le processus d'insertion, le mode optimal est changé par le mode sous-optimal parmi les modes pairs ou impairs selon les règles de mise en correspondance et le bloc modifié est ré-encodé.

La sécurité des données secrètes est assurée par le chiffrement et le brouillage d'une part, et par la sélection des positions de blocs qui est contrôlée par une clé privée, d'autre part. Le schéma promet une excellente préservation de la qualité visuelle et une augmentation faible du débit après l'insertion de la marque.

Xu et al. [8] utilisent la même idée de tatouage basée sur le changement du mode de prédiction (optimal) par le mode sous optimal parmi les modes paires ou impaires selon la parité du bit à insérer. La marque subie un prétraitement qui est un cryptage par une séquence chaotique. Ce schéma applique l'insertion durant le processus d'encodage, il est donc mieux adapté pour les applications en temps réel.

2.3.2.2 Insertion de la marque après compression

Outre les exemples cités précédemment, relatifs à l'insertion de la marque lors de la compression, il existe des schémas qui proposent d'effectuer l'insertion après compression, car la vidéo est stockée et distribuée dans ce format. Pröfrock et al. [39] ont présenté un schéma de tatouage fragile approprié à l'authentification des données. Il se compose d'un transcodeur qui

effectue un décodage entropique, analyse le flux du H.264, calcule la marque qui sera insérée et enfin génère de nouveau le flux H.264. La marque représente la valeur de hachage cryptée et un certificat avec une clé publique. Les LSBs des coefficients DCT luma non codés (*skipped⁹ coefficients*) sont utilisés pour l'insertion. Dans cette approche, la marque peut être supprimée et la vidéo reconstruite.

Iqbal et al. [40] ont proposé une méthode de tatouage fragile pour contrôler l'authentification de la vidéo. La marque utilisée est basée sur les données de hachage de la vidéo. La marque est insérée au cours du processus d'adaptation réalisée à l'aide de GBS (Generic Syntax Description Bitstream), spécifié dans MPEG 21. L'objectif est d'éviter des opérations de cascade de décompression et de re-compression, ce qui pourrait être plus approprié pour des applications en temps réel.

2.3.3 Schémas de tatouage réversible

Le critère de réversibilité est indispensable dans certaines situations qui nécessitent la récupération du document original à partir du document tatoué, particulièrement dans la manipulation de données médicales et militaires.

Toutefois, la majorité des solutions de schémas réversibles qui ont été proposées ont été appliquées sur des images fixes. Nous avons donc introduit une méthode réversible au codec H.264/AVC [41] tel que présenté dans le chapitre 3.

Un autre exemple plus récent est présenté dans [42]. Dans cette méthode les auteurs proposent d'effectuer une insertion uniquement dans les images de type P, dans les IPCM.

Pour augmenter la capacité d'insertion les macroblocs sont forcés à un codage en IPCM, ce qui a pour effet d'augmenter le débit. Par ailleurs, le schéma proposé est fragile et se limite à des applications d'authentification.

Dans cette méthode, la réversibilité est obtenue par la transformation de nombres entiers (la transformée en S) proposée dans [43]. Cette méthode, connue par « Expansion de Différence », est expliquée brièvement à travers l'exemple ci-dessous emprunté de [43]. D'autres méthodes sont présentées dans le chapitre 3.

Exemple :

Soit x et y des valeurs de luminance d'un pixel, $x=206$ et $y=201$.

Pour insérer un bit 1 nous calculons d'abord la moyenne L et la différence H de x et y comme suit :

⁹ Le mode Skip est utilisé avec les images P et B. Dans ce mode aucune information du macrobloc n'est transmise au décodeur. Ce dernier estime le vecteur de mouvement du macrobloc Skip à partir des macroblocs voisins.

$$L = \lfloor \frac{206+201}{2} \rfloor = \lfloor \frac{407}{2} \rfloor = 203$$

$$H = 206 - 201 = 5$$

Le symbole $\lfloor z \rfloor$ représente la fonction floor qui signifie le plus grand entier inférieur ou égal à z .

Une représentation binaire de H donne : $H=101_2$

Pour insérer un bit ($b=1$), b est concaténé à H de manière à obtenir la nouvelle différence $H'=101b_2=1011_2=11$

Nous pouvons calculer les nouvelles valeurs de x et y :

$$x' = 203 + \lfloor \frac{11+1}{2} \rfloor = 209, \quad y' = 203 - \lfloor \frac{11}{2} \rfloor = 198,$$

Le bit secret est extrait à partir du LSB de la différence H' : $b = \text{LSB}(H')$.

2.3.4 Le prétraitement de la marque

La marque subit généralement un prétraitement pour une meilleure adaptation à la séquence vidéo. Dans le cas où la marque serait une image, des transformations pourraient être effectuées pour réduire la quantité d'information à insérer (transformation en niveaux de gris, binarisation, etc.).

En outre, la sécurité de la marque peut être améliorée avant l'étape d'insertion en effectuant un cryptage ou un brouillage des données par exemple. Les fonctions de hachage sont également utilisées en particulier pour les schémas d'authentification où la valeur de hachage de l'image (l'empreinte) serait une partie de la marque.

Un exemple est donné dans [44] où nous proposons une méthode originale de traitement de la marque avant de l'insérer dans les coefficients DCT du codec H.264. Un seul coefficient par bloc luma 4×4 est utilisé pour insérer un bit par la méthode des LSBs. Le traitement consiste à effectuer une rotation de la séquence binaire représentant la marque pour rechercher le meilleur emplacement qui préserve la qualité visuelle de la qualité. Les deux séquences, représentant respectivement la marque et les coefficients DCT utilisés pour l'insertion, sont remplacées temporairement par deux séquences binaires constituées de -1 et 1. Un calcul d'inter-corrélation est ensuite effectué pour trouver le meilleur emplacement de la marque.

Cet emplacement permet d'avoir le maximum de bits 0 de la marque correspondant lors de l'insertion aux coefficients DCT pairs et les bits 1 aux coefficients impairs. La modification des coefficients est ainsi réduite et la qualité visuelle améliorée.

2.4 Evaluation de la qualité vidéo

2.4.1 Notion d'artéfacts

Par rapport à l'information visuelle d'origine, l'image ou la vidéo contient des artéfacts dus aux processus d'acquisition, de traitement, de transmission ou d'affichage. L'artéfact peut être défini comme toute différence visible résultant des limitations techniques lors du processus de communication [45].

Yuen et al.[46] ont présenté un état de l'art sur les distorsions introduites par les techniques de compression vidéo. Ces effets peuvent être accentués ou atténués par l'insertion d'une marque dans le codec vidéo. Nous présentons ci-dessous quelques uns de ces artéfacts.

L'effet de bloc: C'est généralement l'artéfact d'image ou de vidéo les plus visibles. Il est causé par les techniques de traitement de base sur le découpage en blocs, en l'occurrence le processus de codage vidéo où le traitement des blocs se fait indépendamment les uns des autres ce qui réduit la corrélation entre les blocs adjacents et crée ainsi une discontinuité qui fait apparaître les bords.

L'effet de flou: Le flou est un autre artéfact résultant de l'absence de hautes fréquences (qui représentent les détails). Le flou se traduit par une perte des détails spatiaux et la réduction de l'acuité des bords.

Des exemples d'effets de bloc et de flou sont présentés à la figure 2.5.

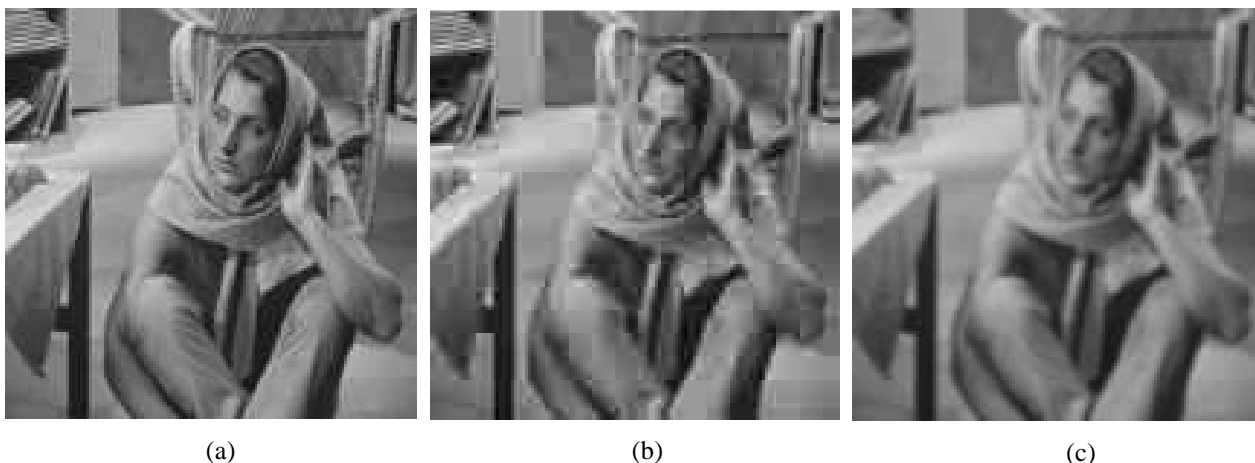


Figure 2.5: Exemples d'artéfacts : (a) Image originale de Barbara (b) L'effet de bloc, (c) Le flou.

L'effet d'ondulation: L'effet d'ondulation est causé par la quantification ou la troncature des coefficients de hautes fréquences (Phénomène de Gibbs) et peut également provenir d'opérations de restauration incorrecte d'image. Ce phénomène apparaît comme des oscillations au niveau des contours.

L'effet escalier : C'est un effet qui apparaît dans les contours obliques qui sont représentés, à la suite du traitement par bloc, comme une succession de contours horizontaux et verticaux formant un escalier.

L'effet de motif de la fonction DCT : Le contenu d'un bloc est représenté par une combinaison de motifs après un traitement par la DCT. Après la quantification, il est possible qu'il ne reste plus que peu de coefficients AC. Les motifs correspondants apparaîtront donc sur l'image.

Les faux bords : ce sont une conséquence de transfert des bords de blocs en raison de la discontinuité de l'effet de bloc à partir d'images de référence dans le cadre prévu par la compensation de mouvement.

2.4.2 Evaluation subjective de la qualité

Pour l'estimation de la qualité visuelle subjective, nous retrouvons plusieurs procédures de test dans les recommandations BT.500-11 de l'union internationale des télécommunications (UIT) [47]. Parmi les méthodes décrites nous donnons un bref aperçu sur la procédure d'échelle de dégradation sur stimulus- double appelée DSIS (*The double-stimulus impairment scale*).

Pour effectuer le test subjectif, il est recommandé que le nombre d'observateurs soit supérieur à 15. Il s'agit de personnes non expertes dans le domaine. Ces personnes doivent avoir une vision normale pour mener à bien la tâche d'évaluation et ne pas errer les résultats.

Les instructions données à l'observateur doivent indiquer ce qui doit être évalué, la procédure expérimentale adoptée ainsi que toutes les informations liées à l'évaluation telles que l'échelle de notation, le nombre et la durée des séquences.

Dans les tests DSIS, l'observateur visualise la version originale de la séquence vidéo suivie de la version dégradée. Il doit juger la dégradation existant entre la vidéo traitée (compressée ou tatouée) et la vidéo de référence. L'échelle utilisée pour juger la qualité est :

- (5) imperceptible ; (4) perceptible mais pas gênante ; (3) un peu gênante ; (2) gênante ;
- (1) très gênante.

La session de test devrait durer environ une demi heure (avec y compris les explications). Environ les cinq premières présentations ne sont pas prises en considération dans l'évaluation car elles permettent à l'observateur de stabiliser son avis.

La session de test englobe un certain nombre de présentations, chaque présentation est composée d'une séquence de référence et de la séquence dégradée. Chaque présentation peut apparaître pour l'évaluation une seule fois (variante 1), ou bien deux fois (variante 2), tels qu'illustré à la figure 2.6.

Après la visualisation des deux séquences, l'observateur donne son appréciation selon l'échelle indiquée précédemment durant un temps T4.

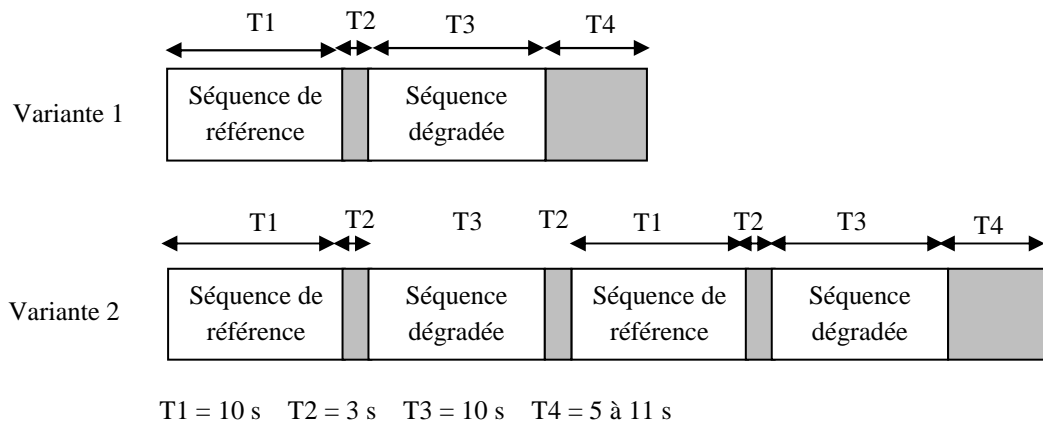


Figure 2.6 : Structure de test des séquences vidéo.

2.4.3 Les métriques de qualité

La mesure en temps réel de la qualité vidéo est une question difficile qui comprend le traitement haute vitesse et une méthode efficace pour quantifier la qualité de la vidéo. De nombreux chercheurs ont travaillé sur la qualité vidéo et la comparaison des différentes méthodes d'évaluation objective [48].

Winkler [49] a présenté dans sa thèse une classification des métriques comme suit :

Les métriques basées sur une comparaison de pixel

La métrique de qualité la plus utilisée est l'erreur quadratique moyenne (MSE). Elle exige un alignement de deux séquences, de telle sorte que chaque pixel de l'image traitée peut être comparé avec les pixels de l'image de référence. Sa formule est donnée dans (2.1).

$$MSE = \frac{1}{m \times n} \sum_{x=1}^m \sum_{y=1}^n (Y_r(x, y) - Y_d(x, y))^2 \quad (2.1)$$

où $Y_r(x, y)$ et $Y_d(x, y)$ sont les niveaux de luminance du pixel (x, y) dans l'image de référence (d'origine) et l'image déformée (traitées) de taille $(m \times n)$ respectivement.

Les modèles "mono-canal" et "multi-canaux"

Les modèles "mono-canal" considèrent le système visuel humain comme un seul filtre spatial, dont les caractéristiques sont définies par la fonction de sensibilité du contraste (CSF). La sortie du système représente la version filtrée du stimulus d'entrée dont la détectabilité dépend d'un seuil.

Dans les modèles multi-canaux, il est supposé que chaque bande de fréquences spatiales est traitée par un canal indépendant. La fonction CSF est juste l'enveloppe canaux de cette sensibilité.

La détection se produit de façon indépendante dans chaque canal lorsque le signal dans cette bande atteint un seuil critère.

Les métriques spécialisées

Ce sont des métriques qui exploitent les connaissances a priori que l'algorithme de transformation subie, tels la compression et le tatouage, et aussi des types d'artéfact et ne nécessitent pas une connaissance approfondie du système visuel humain.

Par ailleurs, Le groupe VQEG [50] a classé les modèles de qualité vidéo en trois catégories différentes en fonction des méthodes de comparaison de la vidéo transformée (traitée) par rapport à l'originale. Nous distinguons :

- des modèles à référence complète (FR, pour *Full reference*) : le signal d'origine doit être présent à l'utilisateur final pour mesurer la qualité. L'erreur quadratique moyenne est considérée comme étant un modèle à référence complète. Le PSNR (*Peak Signal to Noise Ratio*) peut être calculé à partir de MSE comme suit :

$$\text{PSNR} = 10 \times \log_{10} \left(\frac{255^2}{\text{MSE}} \right) \text{ dB} \quad (2.2)$$

Il est clair que la fonction MSE néglige les propriétés du système visuel humain, et n'est pas toujours fiable, c'est la raison pour laquelle de nouveaux modèles sont proposés tel que le modèle présenté par Tan et al. [51][52] qui effectue une détection de l'effet de bloc et du flou pour la vidéo MPEG (voir chapitre 3, sous-section 3.3.2) ;

- des modèles à référence réduite (RR, pour *Reduced reference*) : certaines caractéristiques extraites du signal d'origine sont transmises au niveau du récepteur pour être utilisées dans la mesure de la qualité, tel que le modèle de Choi et al. [53] qui est basé sur la différence quadratique moyenne le long des limites des blocs.
- des modèles sans référence (NR, pour *Non-reference*) : ne nécessite aucune information du signal original au niveau du récepteur, tel que le modèle de Liu et al. [54] qui exploite une structure spécifique d'information (texture et le masquage de luminance).

2.5 Conclusion

Dans ce chapitre nous avons présenté un aperçu de la littérature des travaux relatifs aux techniques de tatouage d'une manière générale, et aux techniques de tatouage appliquées à la vidéo et particulièrement au codec vidéo H.264/AVC.

Malgré la complexité de ce codec, plusieurs méthodes de tatouage ont été proposées en exploitant ses différents modules. Toutefois leur domaine d'application reste limité étant donné la difficulté d'atteindre un bon compromis entre les différents critères. Par ailleurs, nous constatons que très peu de méthodes de tatouage réversible ont été proposées. Dans le prochain chapitre, nous introduisons au codec H.264/AVC une nouvelle méthode de tatouage réversible empruntée des méthodes de tatouage d'image. Son applicabilité aux images de types I et P est étudiée. Une deuxième méthode originale de tatouage réversible est également proposée. Nous présentons son apport par rapport au compromis entre la qualité visuelle, le débit et la capacité d'insertion.

3.1 Introduction

La conception d'un système de tatouage de documents multimédias doit satisfaire certains critères tels que la fiabilité du document, la capacité d'insertion et la réversibilité, afin de répondre aux besoins de l'application. Le tatouage réversible permet la restauration du document initial à partir d'un document tatoué. La réversibilité attire de plus en plus d'intérêt, particulièrement dans certains domaines tels que le domaine médical et militaire où il serait utile d'insérer un tatouage réversible dans l'image entière ou uniquement dans une zone d'intérêt [55][56].

Dans ce chapitre, nous introduisons et adaptons des méthodes de tatouage réversible au codec H.264/AVC. La première méthode est basée sur la technique de l'histogramme décalé, initialement proposée pour les images. Dans ce travail, nous comparons l'applicabilité de cette méthode au niveau des images de types I et P en termes de débit, de capacité d'insertion et de qualité visuelle. Pour cette dernière nous utilisons, en plus du PSNR, une métrique objective qui prend en considération les propriétés du système visuel humain.

Une deuxième méthode consiste à introduire au codec H.264/AVC une approche réversible appliquée initialement aux images JPEG. Pour le maintien de la qualité visuelle et du débit, nous proposons une adaptation liée à la zone d'insertion et à l'élimination des cas d'ambiguïté qui pourraient se présenter à la détection [57].

3.2 Le tatouage réversible

Parmi les méthodes de tatouage réversible existantes on distingue celles qui considèrent que l'imperceptibilité de la marque n'est pas très importante, et il suffit de disposer d'un document multimédia reconnaissable puisqu'il est possible de récupérer l'image ou la vidéo originale [58]. Cela faciliterait l'augmentation relative de la capacité d'insertion. Cependant, pour d'autres auteurs [59], l'imperceptibilité de la marque est considérée comme fondamentale pour permettre à l'utilisateur final d'opérer sur le document marqué, et éventuellement décider de recourir à la version non corrompue dans un second temps si nécessaire. Ainsi, le maintien de la qualité de la vidéo est recherché malgré la disponibilité de la réversibilité.

Un des schémas intuitifs de tatouage réversible appliqué aux images, serait de compresser les données de restauration et de les insérer avec les données secrètes dans les LSBs de l'image originale. Toutefois, la robustesse de ce type de schéma est faible parce que les données intégrées peuvent facilement être perdues [60].

La technique de l'histogramme décalé (SH) a été proposé par Ni et al.[61] pour les images fixes. La technique consiste à sélectionner dans l'histogramme de l'image les paires de niveaux crête/

zéro (c , z) correspondant aux niveaux de gris pour lesquels la fréquence est la plus élevée, respectivement la plus basse. Ensuite, appliquer un décalage entre ces deux niveaux de pixels (c , z) de sorte qu'un espace est créé en c . L'objectif est de maximiser la capacité d'insertion en exploitant la plus grande fréquence pour l'insertion de la marque. Nous introduisons cette méthode au codec H.264/AVC comme expliqué dans la section suivante.

Dans [43], les auteurs proposent une méthode de tatouage réversible à grande capacité pour les images en niveaux de gris. Cette méthode, appelée expansion de différence (DE), est basée sur la transformée en ondelettes de Haar entière, aussi appelée la transformée en S. La valeur de la différence de deux pixels est élargie pour insérer un bit secret. Une autre transformation de nombres entiers est proposée par Coltuc et al. [62] afin de fournir un schéma réversible de haute capacité. Cette méthode est appelée RCM (*Reversible Contrast Mapping*) car elle apporte une modification dans le contraste de l'image. Un bit secret est inséré dans la paire de pixels transformée. Certaines contraintes sont imposées dans ces deux dernières méthodes pour éviter les problèmes de débordement de niveaux. La méthode proposée dans [43] a été exploitée pour insérer d'une manière réversible des logos visibles au H.264/AVC [63].

Dans [64], les auteurs ont proposé un schéma pour insérer des données secrètes dans une image compressée (JPEG) au niveau des blocs 8x8 de coefficients DCT quantifiés, et restaurer les coefficients initiaux après l'extraction de la marque. Dans [65], cette méthode a été appliquée pour le codec MPEG 4 en adoptant quelques modifications pour améliorer la qualité visuelle. Dans ce chapitre une adaptation de cette méthode est présentée pour le codec H.264/AVC dans la section 3.5.

3.3 Méthode réversible basée sur le décalage de l'histogramme

3.3.1 La méthode de Ni

La méthode de tatouage basée sur la technique du décalage de l'histogramme consiste à choisir une paire de l'histogramme de l'image et à appliquer un décalage entre les niveaux de pixels c et z correspondant respectivement au niveau ayant la fréquence maximale ou minimale, créant ainsi un écart au niveau de c [61].

La méthode d'insertion est la suivante :

L étant la luminance du pixel (initialement $L=c$) et w est le bit à insérer,

1^{er} cas: $c > z$

Après le décalage de l'histogramme : $L=c-1$,

Si ($w= 1$) alors $L=L+1=c$, sinon, $L= c-1$ (non modifié),

2^{ème} cas : $c < z$,

Après le décalage de l'histogramme : $L=c+1$,

Si ($w=1$) alors $L=L-1=c$, sinon, $L=c+1$ (non modifié).

Ce tatouage de données est réversible et la détection se fait par une lecture des valeurs de luminance de l'image. Si une valeur de luminance du pixel est égale à p le bit inséré est "1", et si elle est égale à $c-1$ (pour le premier cas) ou $c+1$ (pour le second cas), le bit inséré est "0".

Plusieurs paires peuvent être choisies dans l'histogramme de l'image afin d'augmenter la capacité d'insertion. Les valeurs (c, z) doivent être identifiées comme des informations supplémentaires nécessaires pour la réversibilité du tatouage et doivent être transmises au récepteur [66].

Dans la figure 3.1 l'histogramme de l'image Barbara est représenté. Dans cet exemple, il est possible de choisir $c=159$ correspondant à la fréquence 2250 et $z=15$ correspondant à la fréquence nulle.

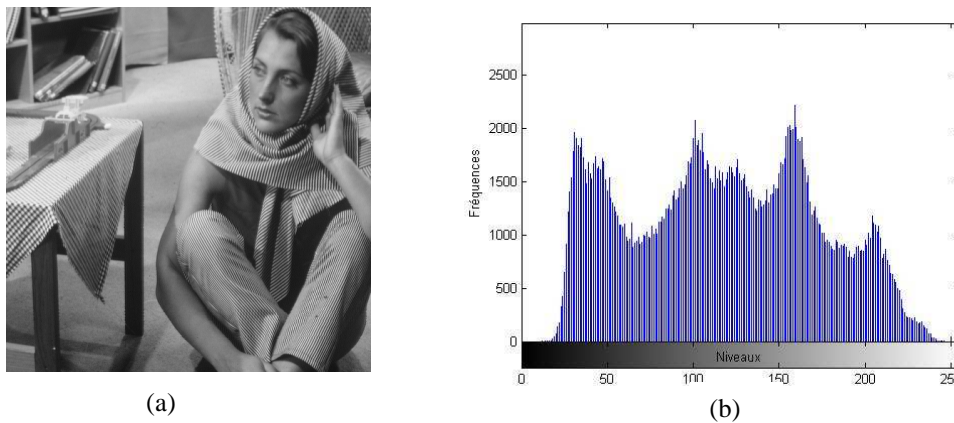


Figure 3.1 : (a) L'image de Barbara (b) L'histogramme de Barbara.

3.3.2 Approche proposée

3.3.2.1 Insertion et extraction de la marque

L'analyse de quelques histogrammes des coefficients DCT quantifiés de séquences vidéo codées avec le H.264/AVC, révèle une forme similaire à la distribution de la Laplace, tel que représenté sur l'exemple de la figure 3.2.

Nous appliquons la technique de l'histogramme décalé (SH) aux coefficients quantifiés DCT luma 4x4 des images de types I et P du codec vidéo H.264/AVC. L'histogramme des coefficients DCT significatifs (non nuls) montre deux côtés de valeurs positives et négatives presque symétriques.

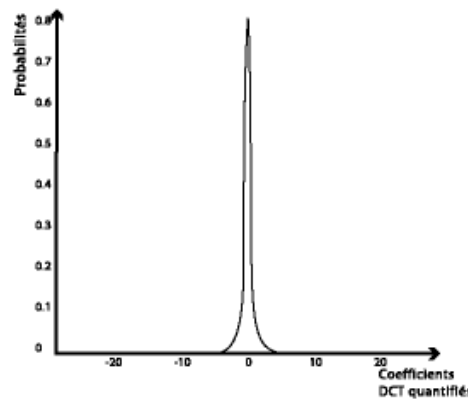


Figure 3.2 : Forme de l'histogramme des coefficients DCT quantifiés pour la séquence Coast-Guard.

Les étapes suivantes décrivent une façon possible pour appliquer la méthode sur le côté positif de l'histogramme:

- Trouver une paire de maximum et minimum (c , z) dans les niveaux positifs significatifs, et décaler les valeurs dans l'intervalle $[c, z-1]$, c'est-à-dire vers la droite où se trouve la fréquence minimale, par incrémentation des niveaux par 1.
- Si le bit de la marque est égal à "1", réduire le coefficient $(c + 1)$ par 1, sinon le coefficient n'est pas modifié.

Pour appliquer la méthode dans les deux côtés (positif et négatif) les étapes précédentes sont répétées dans le côté négatif de l'histogramme en déplaçant les niveaux dans l'intervalle $[z'+1, c']$ vers la gauche. (c' , z') est la paire de valeurs de maxima et minima dans le côté négatif de l'histogramme. Dans notre cas, nous avons appliqué l'insertion dans les deux côtés afin de maximiser la capacité d'insertion et préserver le débit [41].

L'étape de détection de la marque est faite lors du décodage vidéo, après l'étape de décodage entropique. Si la valeur du niveau est égal à c ou c' , cela signifie que le bit de la marque est "1", et si la valeur du niveau est $(c-1)$ ou $(c'+1)$, le bit inséré est égal à "0" comme mentionné précédemment.

3.3.2.2 Evaluation de la qualité de la vidéo

Afin d'évaluer la qualité de la vidéo après insertion de la marque, la qualité objective est mesurée par le PSNR des séquences. Toutefois, lorsque les résultats donnés par le PSNR ne traduisent pas l'estimation subjective de la qualité, il est préférable d'utiliser une deuxième métrique de qualité. Dans notre cas nous utilisons une métrique de qualité basée sur les détecteurs des effets de bloc et de flou [51]. Dans cette métrique, la FFT dans une fenêtre de 32×32 pixels des gradients de l'image de référence et de celles traitées sont calculées. L'énergie ajoutée à la phase des

harmoniques par rapport à l'image de référence est une indication de l'effet de bloc et la perte d'énergie dans l'amplitude des harmoniques serait le flou [52]. Par conséquent, en mesurant les effets de bloc et de flou, il serait possible de juger la qualité vidéo. Les résultats de mesure sont normalisés dans l'intervalle 0 à 100%. Des scores supérieurs à 60% seraient interprétés comme une bonne qualité subjective en termes d'effets de bloc ou de flou.

3.3.2.3 Résultats et analyses

Pour les tests, les 50 premières images des quatre clips vidéo CIF (352 x 288) : Walk, Coast-Guard, Silent et Foreman (avec le logo Siemens) ont été codées avec le paramètre de quantification ($QP = 28$), à une cadence de 30 images/seconde et générant environ 253 à 1139 kbits / seconde. Avec le profil de base, seule la première image a été codée en intra (I) et les images restantes ont été codées en inter (P). Trois tests ont été effectués : La marque a été insérée d'abord dans I, puis dans les 10 premières images P et enfin l'insertion est appliquée dans les deux types d'images I et P. L'objectif est d'estimer la capacité d'insertion des séquences CIF et d'évaluer l'impact de l'insertion de la marque sur la qualité vidéo et l'augmentation du débit.

L'insertion a été effectuée dans les blocs prédits 4x4 étant donné que les blocs 16x16 représentent les zones homogènes dans lesquelles la moindre distorsion pourrait être visible.

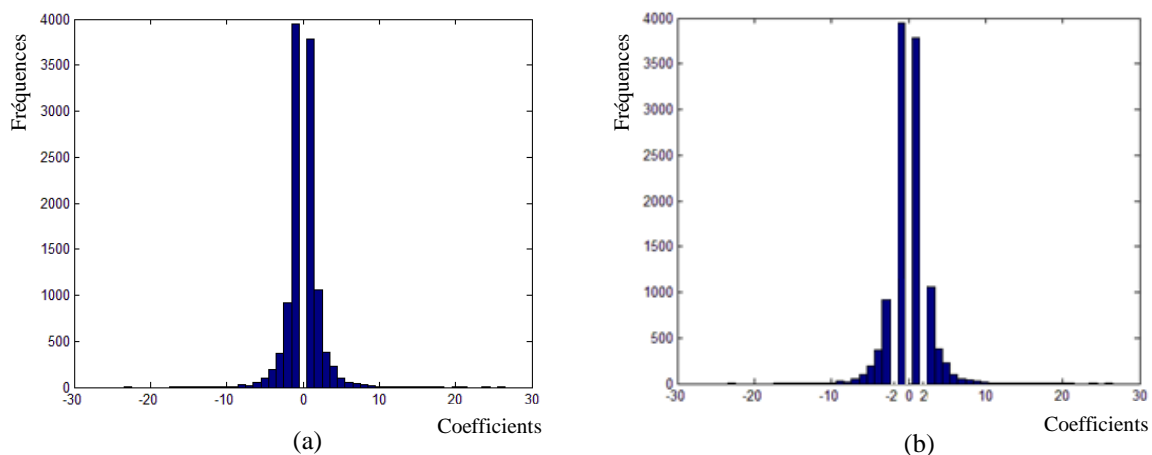


Figure 3.3 : Histogramme des coefficients DCT significatifs de l'image I de la séquence Walk : (a) Avant décalage, (b) Après décalage.

L'insertion de la marque dans les coefficients +1 et -1, qui offrent une capacité d'insertion maximale (voir histogramme, figure 3.3), engendre des artéfacts visibles et une augmentation relativement importante du débit. Cela est dû au fait que les niveaux ± 1 jouent un rôle crucial dans la chaîne de codage du H.264/AVC [41]. Nous avons donc opté pour l'utilisation des coefficients ± 2 pour insérer la marque, car ils semblent offrir un meilleur compromis entre la capacité d'insertion, l'augmentation du débit et la qualité de la vidéo.

a) Le taux d'insertion et la qualité de la vidéo

La capacité d'insertion dépend du contenu de l'image. Pour I, elle varie entre 1123 et 2925 bits correspondant aux séquences Foreman et Coast-Guard, respectivement. Les figures 3.4, 3.5 et 3.6 montrent les quatre images I non marquées et les images marquées correspondantes.

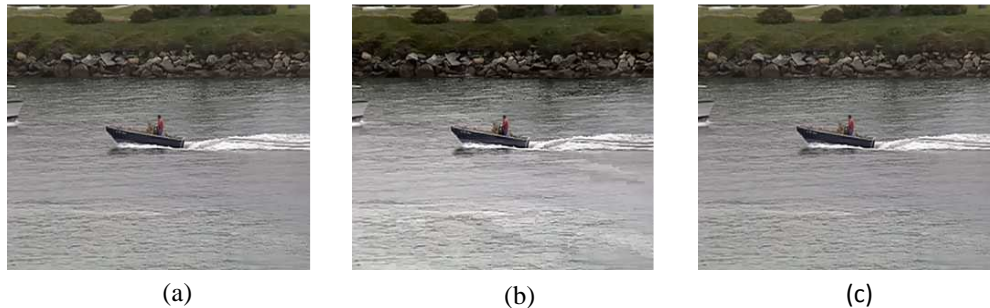


Figure 3.4 : Séquence Coast-Guard (a) Image I non marquée, (b) Image I marquée, (c) Image P marquée.



Figure 3.5 : (a1), (a2) images I des séquences non marquées Walk et Silent respectivement, (b1) et (b2) sont les images marquées correspondantes.



Figure 3.6 : Séquence Foreman : (a) Image I non marquée, (b) Image I marquée, (c) Image P marquée.

Pour ce premier cas (marque insérée dans I), la figure 3.7 montre que l'insertion de la marque par la méthode de l'histogramme décalé produit une distorsion importante. La qualité vidéo objective, pour toutes les images, marquées ou non, chute d'environ 12 à 13.5 dB. Ces valeurs ainsi que la capacité d'insertion et l'augmentation du débit sont enregistrés dans les tableaux 3.1, 3.2, 3.3 et 3.4.

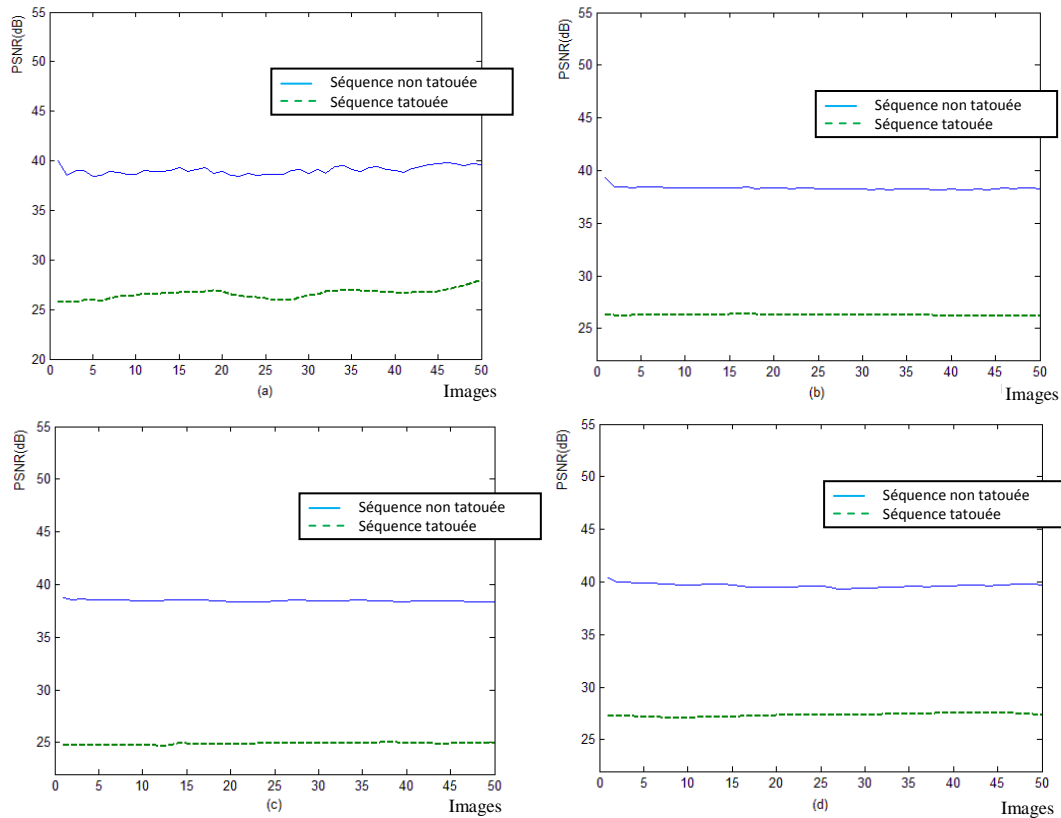


Figure 3.7: PSNR des séquences vidéo marquées et non marquées: (a) Walk, (b) Coast-Guard, (c) Silent, et (d) Foreman.

Les tableaux 3.1, 3.2, 3.3 et 3.4 montrent également l'impact de la marque sur la qualité subjective de la vidéo, sur différentes images (de l'Image 1, F1 à l'image 30, F30). Nous donnons les résultats de qualité de l'image I, de la première image P et des trois autres images P. Bien que la mesure objective (PSNR) indique une importante dégradation, la dégradation apparaîtrait (subjectivement) uniquement dans la première image (I). La plus forte dégradation est mesurée pour la séquence Coast-Guard qui présente un effet de bloc de 72% en raison de la forte capacité d'insertion qu'elle offre (supérieure à la capacité d'insertion de Foreman de plus de deux fois et demi). L'impact de l'insertion de la marque sur le reste des images est négligeable en termes d'effet de bloc et de flou.

Cela peut s'expliquer par le fait que dans la technique de l'histogramme décalé le passage est généralement appliqué à un nombre relativement élevé de coefficients. L'effet de ce changement est transmis avec la prédiction inter aux images P réduisant ainsi le PSNR pour la séquence

entière même dans les images où la marque n'a pas été insérée. Subjectivement, nous pouvons remarquer un changement dans la luminance des images, mais toute éventuelle dégradation causée par l'insertion de la marque reste masquée dans la texture de l'image.

Pour le cas de l'insertion de la marque dans les images P, au lieu d'insérer la marque dans une image, elle est répartie sur plusieurs images étant donné que la capacité d'insertion dans chaque image P est très faible comparée à l'image I. Les tableaux des résultats montrent que la capacité d'insertion pour les quatre séquences varie entre 168 et 5403 bits/10 images P.

Les plus grandes capacités sont observées pour les séquences Walk et Coast-Guard qui présentent également le plus hauts débit (1111.15 et 1139 kbits /seconde, respectivement). La réduction du PSNR est relativement faible, elle varie en conformité avec la capacité d'insertion comprises entre 0.02 et 1.28 dB correspondant aux séquences Silent et Walk respectivement.

En ce qui concerne la qualité subjective, elle est parfaitement préservée, la dégradation est plus élevée pour la séquence Walk qui présente un effet de bloc négligeable de 95%, si nous ne considérons pas l'effet de bloc existant déjà dans les séquences sans marque. En plus de cela, dans notre test, l'insertion de la marque a été effectuée seulement dans 10 images P, le reste des images P du GOP peuvent également être exploitée afin de cacher plus d'informations sans perturber la qualité vidéo.

Tableau 3.1. La capacité d'insertion, l'augmentation de débit et la qualité vidéo de la séquence Walk.

Seq. vidéo Walk	Capacité (bits / im.)	Débit (%)	Diff. PSNR (dB)	L'effet de bloc %					Le flou %				
				F ₁	F ₂	F ₁₀	F ₂₀	F ₃₀	F ₁	F ₂	F ₁₀	F ₂₀	F ₃₀
Non marquée	-	-	-	92	100	100	100	100	100	100	98	100	99
Image I	1975	0.17	12.47	85	95	100	100	100	100	100	98	97	96
Images P	5403	0.42	1.28	92	95	98	99	100	100	100	100	100	100
I & P	7378	0.59	12.38	85	93	99	100	100	100	100	100	100	99

Tableau 3.2. La capacité d'insertion, l'augmentation de débit et la qualité vidéo de la séquence CG.

Seq. vidéo Coast-Guard	Capacité (bits / im.)	Débit (%)	Diff. PSNR (dB)	L'effet de bloc %					Le flou %				
				F ₁	F ₂	F ₁₀	F ₂₀	F ₃₀	F ₁	F ₂	F ₁₀	F ₂₀	F ₃₀
Non marquée	-	-	-	100	100	100	100	98	96	98	99	95	98
Image I	2925	0.19	12.04	72	78	86	90	88	100	100	100	100	100
Images P	4327	0.21	0.81	100	100	100	100	98	98	100	100	98	99
I & P	7252	0.41	12	72	85	83	87	86	100	100	100	100	100

Tableau 3.3. La capacité d'insertion, l'augmentation de débit et la qualité vidéo de la séquence Silent.

Seq. vidéo Silent	Capacité (bits / im.)	Débit (%)	Diff. PSNR (dB)	L'effet de bloc %					Le flou %				
				F ₁	F ₂	F ₁₀	F ₂₀	F ₃₀	F ₁	F ₂	F ₁₀	F ₂₀	F ₃₀
Non marquée	-	-	-	96	95	100	96	100	99	100	99	98	98
Image I	1791	0.56	13.5	89	90	93	91	94	100	100	98	97	97
Images P	168	0.04	0.02	96	95	99	96	100	99	100	99	98	98
I & P	1959	0.60	13.5	89	90	92	93	97	99	100	98	97	96

Tableau 3.4. La capacité d'insertion, l'augmentation de débit et la qualité de la vidéo de la séquence Foreman.

Seq. Vidéo Foreman	Capacité (bits / im.)	Débit (%)	Diff. PSNR (dB)	L'effet de bloc %					Le flou %				
				F ₁	F ₂	F ₁₀	F ₂₀	F ₃₀	F ₁	F ₂	F ₁₀	F ₂₀	F ₃₀
Non Marquée	-	-	-	89	92	100	100	100	100	100	100	100	99
Image I	1123	0.23	12.3	91	95	100	100	100	97	97	97	97	95
Images P	839	0.19	0.41	89	92	100	100	100	100	100	100	100	99
I & P	1962	0.43	12.25	91	95	100	100	100	97	97	97	97	95

En ce qui concerne l'insertion réalisée dans les images I et P, les résultats obtenus pour la qualité de la vidéo sont proches de ceux obtenus pour l'insertion dans I. En outre, on peut remarquer que, l'effet de flou d'image est stable, voire améliorée (diminution du flou) pour les séquences Walk et Coast-Guard. Cela est dû au fait que la marque agit comme un dither dans l'image. En effet, la création de détails artificiels a pour effet de réduire le flou de l'image en particulier dans les zones homogènes.

La capacité d'insertion est évidemment beaucoup plus intéressante car il est s'agit de l'addition des capacités des deux cas précédents.

b) L'Augmentation du débit

Les tableaux 3.1, 3.2, 3.3 et 3.4 montrent que le taux d'augmentation le plus faible et le plus élevé du débit est observé pour la séquence Silent. Le taux le plus faible (0.04%) correspond à l'insertion de la marque dans les images P, et le taux le plus élevé (0.6%) correspondant à l'insertion de la marque dans les deux images I et P. Cependant, pour avoir une meilleure idée sur ce que représente cette augmentation par rapport à la capacité d'insertion, nous avons calculé le taux entre la capacité d'insertion et l'augmentation de débit et nous avons remarqué que pour tous les cas, la capacité d'insertion est plus élevée que l'augmentation de débit introduite par l'insertion de la marque. Le pire des cas est observé lorsque la marque est insérée dans I pour la séquence Walk pour laquelle l'augmentation du débit est 0.17%, équivalente à 1904.64 bits/seconde, pour une capacité d'insertion de 1975 bits, et le meilleur des cas est celui de la séquence Coast-Guard où la marque est insérée dans les images P avec une augmentation de débit de 2990.08 bits / seconde pour une capacité d'insertion de 4327 bits. Ceci est intéressant dans la mesure où au lieu d'envoyer des données dans des canaux différents, elles peuvent être envoyées dans le même canal avec un débit inférieur.

Enfin, le choix de marquer les images I, P ou bien les deux en même temps pour assurer la continuité de la marque dépend des besoins de l'application. Bien sur, dans le cas où la marque apporte une dégradation visible, il est nécessaire de réduire la capacité d'insertion (en réduisant le nombre de blocs marqués dans le macrobloc) ou bien d'opter pour le tatouage des images de type P. En général, selon l'augmentation du débit tolérée et la quantité d'informations dont nous

avons besoin de transporter dans la vidéo, il est possible de choisir entre les images P qui préservent le mieux la qualité vidéo, ou les images I en raison de leur importance dans la séquence codée, ou dans les deux d'entre eux de sorte que la capacité d'insertion soit maximisée.

3.4 Méthode réversible basée sur les séquences nulles des blocs DCT

3.4.1 Les méthodes de Cheng et Gujjunoori

Dans la méthode proposée par Chang et al. [64] l'insertion des données est effectuée dans les blocs DCT 8x8 des images JPEG. Le processus d'insertion est le suivant:

Plusieurs ensembles R_i ($1 \leq i \leq 9$) sont définis comme illustré à la figure 3.8 (a):

$R_i = \{ r_{i,1}, r_{i,2}, \dots, r_{i,k} \}$, k varie entre 4 et 7 selon l'ensemble R_i . Par exemple, R_1 dans la figure 3.8 (b) est égal à: $R_1 = \{0, 0, 0, 0, 2, 2, 3\}$.

Un bit secret est inséré dans la séquence de zéros successifs de longueur b_i , qui s'étend à partir de la composante de fréquence la plus élevée ($r_{i,1}$) à la composante la plus basse ($r_{i,k}$). Cette séquence de zéro doit contenir au moins deux zéros pour chaque ensemble R_i ($b_i \geq 2, 1 \leq i \leq 9$).

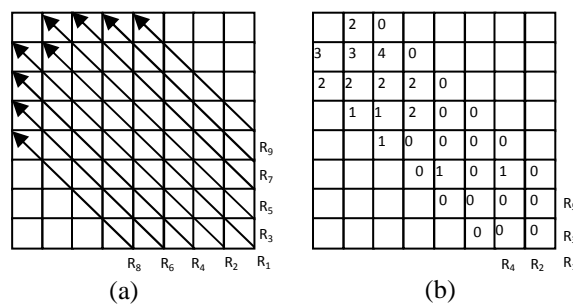


Figure 3.8 : (a) Ensembles d'insertion (b) Exemple de coefficients quantifiés.

$z_{i,1}$ représente la valeur zéro de la fréquence la plus basse d'un ensemble R_i , et $z_{i,2}$ représente le composant inférieur droit de $z_{i,1}$. Dans l'exemple de la figure 3.8 (b), pour l'ensemble R_1 , $z_{1,1} = r_{1,4}$ et $z_{1,2} = r_{1,3}$, $z_{1,2}$ contiendra le bit secret.

Dans cette méthode, les auteurs effectuent l'insertion de plusieurs bits dans le même bloc 8x8. En plus, il est nécessaire d'éliminer les ambiguïtés qui apparaissent à l'étape de la détection en modifiant un second bit dans le même ensemble. Par exemple, le cas de l'ensemble R_2 : $z_{2,1} = r_{2,2} = 0$, et $z_{2,2} = r_{2,1}$, à l'étape de détection $r_{2,3}$ peut être confondu avec $z_{2,2}$ et le bit inséré en tant que 1, donc il doit être incrémenté avant insertion du bit en $z_{2,2}$ pour éliminer cette ambiguïté.

Cette méthode a été adaptée dans [65] pour le codec vidéo MPEG-4. Les différences principales sont les suivantes :

- Dans le but d'améliorer la qualité visuelle, le bit secret est inséré dans une position fixe du coefficient zéro, dans une fréquence $r_{\lceil k/2 \rceil}^{10}$ du milieu de l'ensemble R,
- Deux bits, au lieu d'un, sont insérés dans un coefficient en vue d'améliorer la capacité d'insertion.

Les auteurs suggèrent d'effectuer l'insertion dans les composantes de luminance des images I après la DCT et la quantification. Les blocs 8x8 sont identifiés pour insérer le bit secret dans un coefficient zéro x , $x = r_{i,j}$, $r_{i,j}$ est une fréquence moyenne à partir d'un ensemble R_i qui satisfait la condition $r_{i,j} = 0$ pour $j \leq j'$.

Dans ce schéma, une fonction S est utilisée pour coder deux bits secrets W_1 et W_2 comme suit:

$$S(w_1, w_2) = \begin{cases} 0 & \text{si } w_1 = w_2 = 0, \\ 1 & \text{si } w_1 = 0 \text{ et } w_2 = 1, \\ -1 & \text{si } w_1 = 1 \text{ et } w_2 = 0, \\ 2 & \text{si } w_1 = w_2 = 1, \end{cases} \quad (3.1)$$

Pour insérer une paire de bits, un coefficient x est transformé en x' comme suit:

$$x' = \begin{cases} S(w_1, w_2) & \text{si } x = 0 \\ x+2 & \text{si } x > 0 \\ x-1 & \text{si } x < 0 \end{cases} \quad (3.2)$$

Dans le cas où le coefficient x n'est pas égal à 0, l'insertion n'est pas effectuée mais l'ambiguïté est éliminée en modifiant le coefficient en utilisant (3.2).

Lors de l'étape de détection, il est possible de déduire la valeur des bits secrets selon la valeur extraite x' . Les bits extraits w_1, w_2 sont déduits comme suit:

$$(w_1, w_2) = \begin{cases} (0,0) & \text{si } x' = 0 \\ (0,1) & \text{si } x' = 1 \\ (1,0) & \text{si } x' = -1 \\ (1,1) & \text{si } x' = 2 \end{cases} \quad (3.3)$$

3.4.2 Approche proposée

3.4.2.1 Schéma adapté

Pour adapter la méthode proposée dans [65] pour le codec H.264/AVC, nous proposons d'insérer les données secrètes pendant le processus de compression, dans les coefficients DCT 4x4 quantifiés.

En raison de la taille relativement petite des blocs, un seul ensemble R ($R = \{r_i, i = 1, \dots, 16\}$), composé des coefficients quantifiés dans le balayage en zigzag, est défini comme illustré à la figure 3.9 (a).

¹⁰ $\lceil k/2 \rceil$ représente la fonction « ceil » qui donne le plus petit entier supérieur ou égal à $(k/2)$.

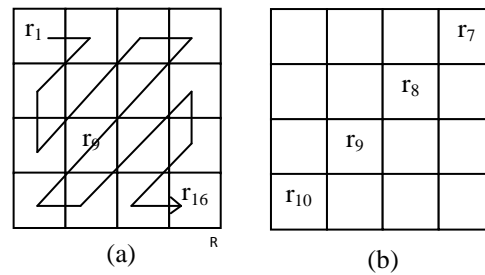


Figure 3.9 : (a) Ensemble R pour l'insertion des données,
(b) Position d'insertion dans la méthode proposée.

Ainsi, selon le schéma de Gujjunoori et al. [65], l'insertion est réalisée dans le coefficient r_9 dans les blocs où les coefficients de r_9 à r_{16} sont nuls. Les équations (3.1) et (3.2) sont utilisées pour insérer deux bits et éliminer l'ambiguïté comme expliqué dans la section 3.4.1.

3.4.2.2 Nouveau schéma

Nous proposons d'apporter des modifications qui visent à améliorer principalement la capacité d'insertion et la qualité de la vidéo. Les processus d'insertion, d'extraction et de restauration sont expliqués dans les sous-sections suivantes.

a) Processus d'insertion

Chaque bloc DCT 4x4 sélectionné pour le processus d'insertion est caractérisé par la séquence de zéro successifs (formée de coefficients r_7 à r_{16}).

Pour augmenter la capacité d'insertion, nous utilisons l'anti-diagonale des coefficients nuls comme illustré à la figure 3.9 (b). L'insertion est réalisée par incrémentation ou décrémentation du coefficient zéro en fonction de la règle de mise en correspondance présentée dans le tableau I, de manière à insérer trois coefficients au lieu de deux.

Dans le cas de la méthode de Gujjunoori et al. [65] l'ambiguïté peut être éliminée par une incrémentation par 2 (voir formule (3.2)) ce qui pourrait altérer la qualité de la vidéo. Dans notre cas, pour éliminer l'ambiguïté qui peut apparaître lors du décodage, dans le cas où un coefficient r_i de l'anti-diagonale est initialement égal à 1 ou -1, le coefficient est incrémenté (ou décrémenté) par 1 selon (3.4) pour préserver la qualité visuelle.

$$\begin{aligned} r_i &= r_i + 1 \text{ si } (r_i > 0 \text{ et } i = 7, 8 \text{ or } 9) \\ r_i &= r_i - 1 \text{ si } (r_i < 0 \text{ et } i = 6, 7, 8 \text{ or } 9) \end{aligned} \quad (3.4)$$

Il est clair à partir de (3.4) que cette modification ne s'applique pas seulement dans le cas où r_i est égal à -1 ou 1 (ce qui permettrait d'extraire correctement les données secrètes), mais aussi pour des valeurs non nulles, pour obtenir le critère de réversibilité expliqué à l'étape d'extraction.

Tableau 3.5. Règles de mise en correspondance pour l'insertion.

Données secrètes	Modification des coefficients
000	Pas de modification
001	$r_8 = 1$
010	$r_9 = 1$
011	$r_{10} = 1$
100	$r_7 = -1$
101	$r_8 = -1$
110	$r_9 = -1$
111	$r_{10} = -1$

b) Processus d'extraction et de restauration des données

A l'étape d'extraction, les données secrètes sont simplement déduites à partir des coefficients de l'anti-diagonales. Si un seul coefficient est différent de zéro et est égal à 1 ou -1 les données secrètes sont déduites à partir du tableau 3.5.

La vidéo originale peut être récupérée par la restauration du coefficient r_i , comme suit:

$$\begin{aligned} r_i &= r_i - 1 \text{ if } (r_i > 1 \text{ and } i = 7, 8 \text{ or } 9) \\ r_i &= r_i + 1 \text{ if } (r_i < -1 \text{ and } i = 6, 7, 8 \text{ or } 9) \end{aligned} \quad (3.5)$$

3.4.2.3 Résultats et analyses

Le schéma de tatouage réversible que nous avons proposé a été implémenté en utilisant le langage de programmation C++. Pour les expérimentations, six séquences vidéo de format CIF (Common Intermediate Format) ont été utilisées: Bridge-Close, Coast-Guard, Hall, Mobile, Mother et Silent. Pour chaque séquence, 30 images ont été codées avec le profil de base, à un paramètre de quantification $Q_p = 28$ et une fréquence d'image de 30 images / seconde. La première image est codée en intra alors que les images restantes sont codées comme des trames P.

Une séquence binaire pseudo-aléatoire a été utilisée en tant que données secrètes. Elle a été d'abord introduite dans l'image I conformément à la méthode décrite précédemment, en intégrant l'algorithme dans le codec H.264. Pour évaluer la qualité de la vidéo, nous utilisons le PSNR.

Dans ce qui suit nous comparons tout d'abord l'utilisation des images P et I, puis nous comparons notre approche à celle proposée dans [65].

a) Insertion de données dans les images de type I et P

L'insertion est réalisée dans l'image I seule, puis dans les 30 images P. Un seul coefficient est modifié dans un macrobloc, afin de maintenir les caractéristiques de la vidéo. Les résultats liés à la capacité d'insertion, du PSNR et l'augmentation de débit sont présentés dans le tableau 3.6.

Les résultats montrent que l'exploitation des images P améliore la capacité d'insertion, c'est parce que les trames P sont nombreuses comparées à I. La qualité de la vidéo est également meilleure en raison de la réduction du nombre de blocs tatoués dans une image P en comparaison à l'image I. Des exemples d'images sont représentés à la figure 3.10 dans laquelle nous pouvons observer la différence de marquage entre les images I et P notamment pour la séquence Hall.

Cependant, contrairement à l'image I pour lesquelles l'augmentation du débit est négligeable, l'insertion des données dans les images P cause une augmentation de débit relativement élevée. Le taux le plus élevé est de 4.39% constaté pour la séquence Coast-Guard. Dans ce cas, le débit de la séquence non marquée étant égale à 1311.65 kbits /seconde, la hausse est de 54.58 kbits/seconde, ce qui est nettement supérieur à la capacité d'insertion (23.34 kbits). Ainsi l'insertion des données dans les images P ne peut pas être appliquée pour économiser la bande passante du réseau, avec une telle configuration du codec.

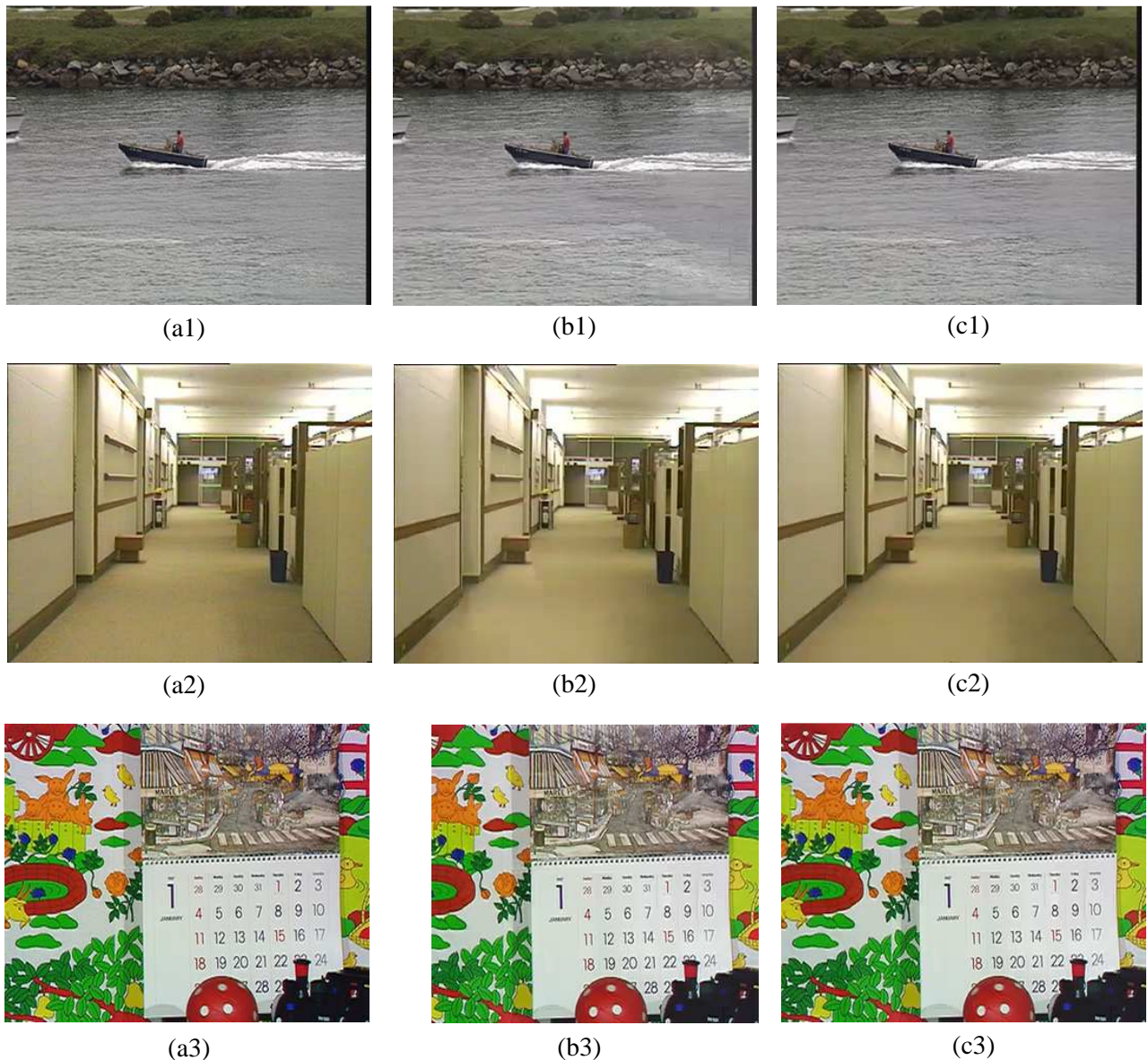


Figure 3.10: (a1), (a2) et (a3) représentent les images I non marquées des séquences Coast-Guard, Hall et Mobile respectivement. (b1), (b2) et (b3) sont les images correspondantes marquées. (c1), (c2) et (c3) sont les premières images P correspondantes marquées.

Tableau 3.6. Capacité d'insertion, PSNR et augmentation du débit pour six séquences vidéo.

Séquences vidéo	Capacité d'insertion (bits)		Réduction de PSNR (dB)		Augmentation du débit (%)	
	Images P	Image I	Images P	Image I	Images P	Image I
Bridge-Close	5882	887	0.20	1.38	3.03	0.40
Coast-Guard	23906	1152	0.87	4.9	4.39	0.22
Hall	3043	645	0.15	4.39	3.41	0.66
Mobile	17454	873	0.33	3.19	3.84	0.08
Mother	1711	730	1.4	3.62	3.97	0.04
Silent	4876	1018	0.13	3.43	4.05	0.88

b) Comparaison entre les méthodes

La méthode décrite dans [65] a été adaptée au codec H264/AVC comme décrit dans la section 3.4.1. La méthode basée sur le décalage de l'histogramme (section 3.3.2) a été implémentée dans les mêmes conditions. L'insertion a été appliquée aux images P. Les résultats sont présentés dans le tableau 3.7 et à la figure 3.11.

Tableau 3.7. Comparaison entre nos approches DCT et SH et celle de Gujjunoori

Séquences vidéo	Réduction du PSNR (dB)			Augmentation du débit (%)		
	Notre approche (DCT)	Notre approche (SH)	Approche de Gujjunoori	Notre approche (DCT)	Notre approche (SH)	Approche Gujjunoori
Bridge-Close	0.20	0.38	0.33	3.03	2.08	2.62
Coast-Guard	0.87	1.6	1.21	4.39	1.9	4.00
Hall	0.15	0.24	0.52	3.41	0.7	3.18
Mobile	0.33	0.97	0.63	3.84	1.5	3.13
Mother	0.06	0.006	1.4	3.97	0.16	3.46
Silent	0.13	0.46	0.30	4.05	1.06	4.06

Comme prévu, même si l'augmentation du débit est relativement élevée à cause de l'utilisation des coefficients nuls, l'application de la règle de mise en correspondance proposée a permis d'améliorer considérablement la capacité d'insertion, principalement pour les séquences à texture hétérogène (riche en détails) telles que la séquence Coast-Guard pour laquelle le taux d'augmentation de la capacité dépasse les autres méthodes d'environ 50% (figure 3.11). En outre, la méthode utilisée pour éliminer l'ambiguïté a très probablement participé à améliorer la qualité de la vidéo, en réduisant les cas de débordement qui peuvent apparaître comme des taches noires (figure 3.12 et figure 3.13). Quant à l'approche SH, elle présente les meilleurs résultats en termes de débit, toutefois sa capacité d'insertion reste tributaire du nombre des images P de la séquence et de la disponibilité du niveau correspondant à la fréquence maximale (dans notre cas ± 2).

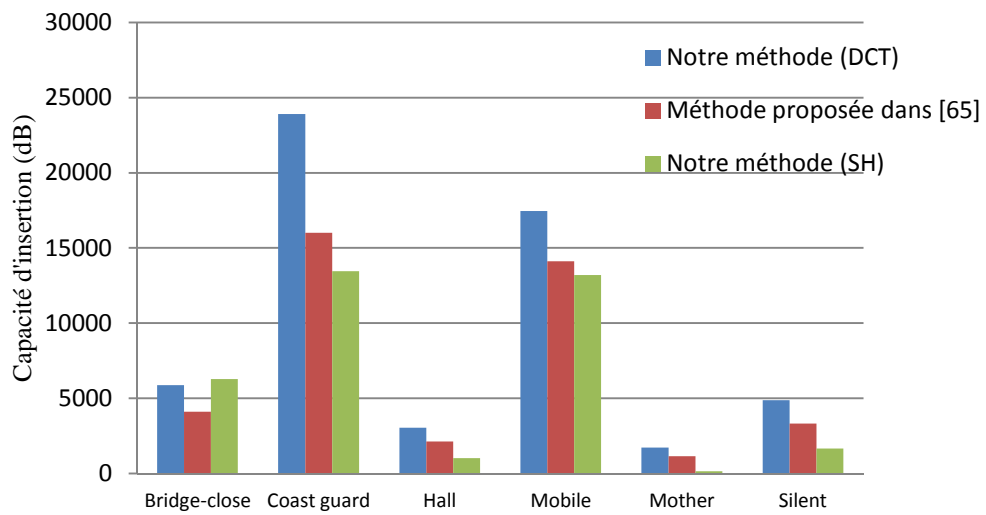


Figure 3.11 : Capacité d’insertion pour six séquences vidéo pour trois méthodes.

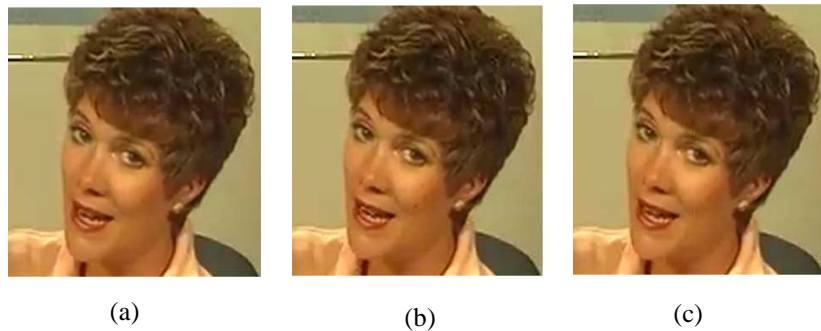


Figure 3.12 : Séquence Mother, 13^{ème} image: (a) Séquence non marquée, (b) Séquence marquée selon Gujjunoori, (c) Séquence marquée selon notre méthode.

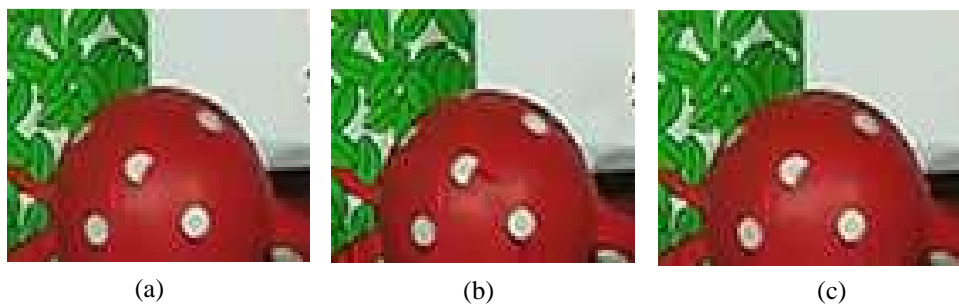


Figure 3.13 : Séquence Mobile, 25^{ème} image: (a) Séquence non marquée, (b) Séquence marquée selon notre méthode (SH), (c) Séquence marquée selon notre méthode (DCT).

3.5 Conclusion

Dans ce chapitre, nous avons proposé deux méthodes de tatouage réversible pour le codec H.264/AVC.

La méthode basée sur le décalage de l'histogramme peut être appliquée à la norme vidéo H264/AVC. En effet, malgré la grande différence entre la qualité objective (PSNR) des séquences vidéo tatouées et non tatouées, la qualité subjective est relativement maintenue notamment pour les images de type P.

Pour les quatre séquences vidéo de test, nous avons remarqué que l'augmentation du débit varie selon le type d'image et le nombre de bits insérés, mais elle ne dépasse jamais la capacité d'insertion. Ceci représente un avantage pour transporter des informations dans la séquence vidéo au lieu d'un second canal.

Pour la deuxième méthode de tatouage réversible basée sur les séquences nulles des coefficients DCT quantifiés, les images de type P peuvent être exploitées pour améliorer la capacité d'insertion et la qualité de la vidéo par rapport aux images I. Ceci est utile pour les applications nécessitant des méthodes réversibles qui préservent la qualité de la vidéo tout en fournissant une capacité d'insertion satisfaisante.

Toutefois, nous remarquons dans ces méthodes pour lesquelles l'insertion se fait dans les blocs DCT 4x4, qu'une restriction du nombre de bits à insérer dans le macrobloc est toujours nécessaire pour maintenir la qualité de la vidéo. Ceci est dû à la propagation d'erreur introduite au niveau du décodage (voir chapitre 5).

Dans le prochain chapitre, nous montrons l'intérêt d'utiliser les modes de prédiction pour l'insertion de la marque afin d'augmenter la capacité d'insertion sans perturber la qualité visuelle.

4.1 Introduction

Dans ce chapitre nous étudions la possibilité d'exploiter les modes de prédiction intra pour l'insertion d'une marque. En effet, l'utilisation de la prédiction intra semble être intéressante dans la mesure où le calcul des résidus se ferait par rapport aux nouveaux modes obtenus, ce qui implique la préservation de la qualité de la vidéo. De ce fait, il serait possible d'obtenir une capacité relativement grande puisque tous les blocs peuvent être concernés par le marquage sans craindre la propagation d'erreur (voir chapitre 5) qui serait minimale, si elle existait.

Comme indiqué dans le chapitre 2, le principe du schéma proposé par Yang et al. [6] et Hu et al. [7] se base sur la modification du mode optimal (le meilleur mode) par le mode sous optimal. Ce schéma nécessite une connaissance préalable des modes de prédiction. Il fut réadapté pour les applications en temps réelles par Xu et al. [8] qui applique la méthode durant le processus de compression en impliquant les blocs 4x4 DCT choisis par une séquence chaotique pour renforcer la sécurité et la robustesse du marquage.

Dans ce chapitre nous présentons une nouvelle approche de l'exploitation des modes de prédiction intra 4x4 pour le tatouage des données [67]. L'idée est basée sur le découpage des modes de prédictions intra en quatre groupes composés de modes qui ont des directions de prédiction proches. L'insertion des données secrètes est basée sur l'application de modifications entre les modes d'un même groupe. Cette approche est basée sur la supposition que l'application d'une modification directe entre les modes de directions proches pourrait:

- préserver les temps d'encodage et de décodage par la réduction du nombre supplémentaire de procédures de calcul générées par la recherche des modes sous-optimaux, comme c'est le cas pour les méthodes proposées précédemment [6][7] ;
- réduire le temps de prétraitement en évitant l'exploitation d'une connaissance préalable des modes utilisés pour la mise en correspondance (*mapping*) [6][7];
- maintenir la qualité visuelle de la séquence vidéo ;
- limiter l'augmentation du débit.

L'augmentation de la capacité d'insertion repose sur le groupe composé de quatre modes, car il permet l'insertion de deux bits par mode au lieu d'un, comme cela est expliqué dans la section 4.7.

4.2 La prédiction Intra-images

Comme indiqué dans le chapitre 1 (section 1.3.3), la norme H.264/AVC utilise la propriété de corrélation spatiale à travers la prédiction Intra. Un bloc (ou macrobloc) peut être prédit à partir des blocs généralement situés au-dessus et à gauche (figure 4.1(a)) car ils ont déjà été codés et

reconstruits. La différence avec le bloc courant est alors codée et représentés avec un nombre réduit de bits par rapport à celle utilisée pour la transformation directe du bloc lui-même.

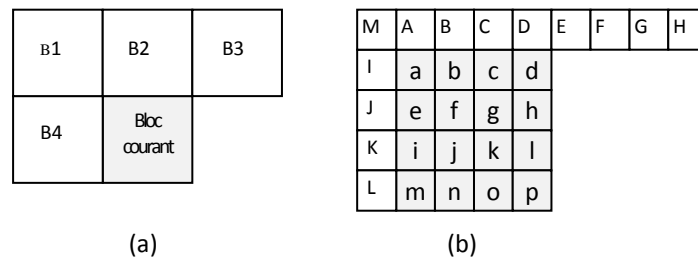


Figure 4.1 : (a) Blocs voisins se positionnant en haut et à gauche du bloc en cours, (b) Echantillons de prédiction.

4.3 Les modes de prédiction intra 4x4

Les modes de prédiction 4x4 conviennent mieux pour les zones riches en détails dans l’image. Le macrobloc est divisé en blocs de luminance de taille 4x4 et chacun des seize échantillons du bloc (de a à p tel qu’illustré à la figure 4.1.(b)), est prédit selon les échantillons précédemment encodés et reconstruits (de A à M). Neuf modes sont définis dans la prédiction 4x4 des blocs de luminance. La figure 4.2 représente les directions des neuf modes de prédiction avec la direction. Dans le mode de prédiction DC la moyenne des valeurs de échantillons qui se situent en haut et à gauche du pixel en cours est utilisée pour prédire le bloc entier [2][3]. Pour ce mode, le codeur affecte une valeur par défaut dans le cas de non disponibilité des échantillons voisins.

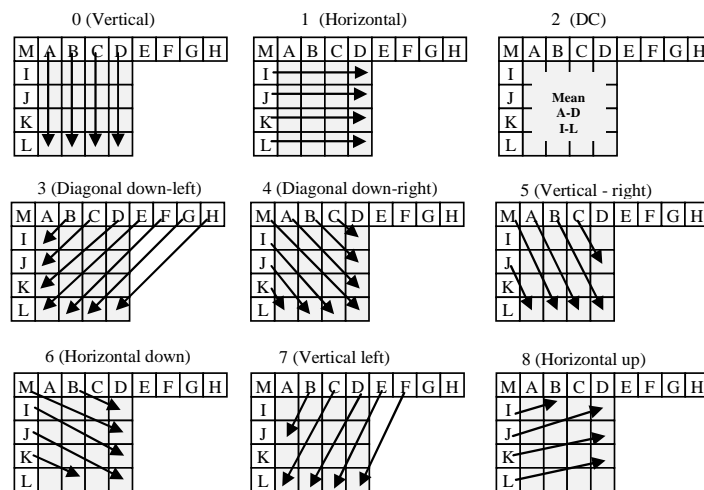


Figure 4.2 : Les neuf directions pour la prédiction Intra.

4.4 Le mode de décision optimal pour la prédiction intra 4x4

La technique d’optimisation du taux de distorsion (RDO) est utilisée pour atteindre la meilleure performance de codage. Pour la prédiction Intra 4x4, tous les modes de prédiction sont testés et le mode de décision optimal est obtenu en minimisant la formule suivante [68].

$$J = D + \lambda_{mode} * R \tag{4.1}$$

D et R représentent respectivement la distorsion et le débit estimé pour le codage du mode. λ_{mode} indique le Lagrangien qui dépend du paramètre de quantification QP. λ_{mode} peut être calculé en appliquant la formule (4.2).

$$\lambda_{\text{mode},I,P} = 0.85 * 2^{(QP-12)/3} \quad (4.2)$$

4.5 Le mode le plus probable

Afin d'effectuer une compression du mode de prédiction appelé le meilleur mode (MM) et représenté par les chiffres de 0 à 8, le codec H.264/AVC définit la notion du mode le plus probable (MPM) qui est une estimation du mode de prédiction en utilisant les blocs se situant en haut et à gauche (B_1 et B_2 dans la figure 4.1. (a)) du bloc en cours. Le MPM correspond à la valeur de prédiction la plus petite parmi les valeurs des modes de B_1 ou B_2 . Si un de ces modes n'est pas disponible, la valeur du MPM est fixée à 2 (mode DC).

Pour chaque bloc le mode est signalé par un paramètre : Si le meilleur mode correspond au MPM, le paramètre P1 est égal à 1, autrement un autre paramètre P2 indique le changement du mode. Si MM est plus petit que MPM le mode de prédiction est ($P2 = MM$), autrement ($P2 = MM - 1$), ainsi les modes sont signalés avec uniquement huit valeurs [3].

4.6 Méthode de marquage basée sur la prédiction intra et les règles de mise en correspondance

Dans cette section nous donnons un aperçu de la méthode proposée par Yang et al. [6]. Cette méthode se base essentiellement sur :

- l'insertion de la marque dans les blocs en changeant le mode optimal (MM) par le mode sous-optimal ;
- l'application des règles de mise en correspondance entre les bits de la marque et les modes de prédiction 4x4.

Dans l'approche présentée dans [6], les blocs sélectionnés pour l'insertion de la marque doivent satisfaire certaines conditions qui ont pour objectif de maintenir le débit. Par exemple, le bloc sélectionné ne doit pas avoir un mode de prédiction égal au MPM.

Chaque trois bloc DCT 4x4 des images intra qui rencontre les conditions d'insertion de la marque constitue un groupe.

Les neuf modes possibles des blocs sont divisés en deux groupes: les modes impairs (1, 3, 5, 7) qui vont correspondre au bit d'information "1" et les modes pairs (0, 2, 4, 6, 8) vont

correspondre au bit "0". Ainsi, nous pouvons extraire les bits d'informations relatives à chaque bloc.

Chaque deux bits de la marque sont modulés aux modes de prédiction de ces trois blocs, et seulement un bloc est nécessaire pour changer le mode de prédiction pour insérer deux bits. Dans le processus d'insertion, le mode optimal est modifié par un mode sous-optimal parmi les modes pairs ou impairs selon les règles de correspondance. Les blocs modifiés sont ensuite ré-encodés.

Le tableau 4.1 présente les règles de correspondance nécessaires pour l'insertion de deux bits 00 de la marque dans un groupe de blocs. Si nous supposons que les modes des trois blocs sont 4,3 et 6, nous faisons d'abord une extraction de l'information correspondante à ces modes, qui est 0, 1 et 0. Dans ce cas nous avons besoin d'effectuer le changement du mode du 2^{ème} bloc pour insérer 00. Dans le cas où il s'agirait d'insérer les bits 10 dans ce même groupe de blocs, on n'a pas besoin d'effectuer de changement. A l'étape de l'extraction, il est possible d'extraire directement deux bits à partir de bits d'information extraits pour chaque groupe de blocs.

Tableau 4.1. Règles de correspondance liées à l'insertion des bits 00 dans un groupe de bloc.

Les bits à insérer	Les bits d'information des blocs			Le bloc qui doit subir une modification
	Q1	Q2	Q3	
00	0	0	0	Pas de changement
	0	0	1	3eme bloc
	0	1	0	2eme bloc
	0	1	1	1er bloc
	1	0	0	1er bloc
	1	0	1	2eme bloc
	1	1	0	3eme bloc
	1	1	1	Pas de changement

La sécurité des données secrètes est assurée par le chiffrement et le brouillage d'une part, et par la sélection des positions de blocs qui est contrôlée par une clé privée, d'autre part. Le schéma promet un PSNR élevé et une augmentation faible du débit après l'insertion de la marque.

Les méthodes de tatouage basées sur les modes de prédictions intra ont pour principal avantage de préserver la qualité de la vidéo, car il est possible de re-encoder le bloc après modification. Cependant, dans la méthode proposée dans [6] l'utilisation des règles de mise en correspondance nécessitent la connaissance préalable des modes de prédiction d'un groupe (c'est-à-dire que l'encodage de ces blocs devra se faire 2 fois : la première pour connaître le mode de prédiction et la 2eme se fait après l'insertion de la marque) ce qui constitue une contrainte pour l'application de la méthode pour les applications en temps réel.

De plus, l'application directe de cette méthode (sans règles de mise en correspondance) tel qu'expliqué dans [8] limiterait le nombre d'insertion à 1 bit par bloc et entrainerait l'augmentation du débit.

Par ailleurs, ces méthodes ont été proposées en considérant l'image comme une seule tranche. L'utilisation de plusieurs tranches (voir groupe de tranche, chapitre 1, figure 1.16) poserait le problème de disponibilité de blocs à modifier car la prédiction ne se fait qu'à l'intérieur d'une seule tranche.

Pour pallier ces problèmes, nous proposons une approche qui pourrait être utilisée dans les applications en temps réel pour augmenter la capacité d'insertion. Dans l'approche proposée, nous modifions le mode optimal par un mode qui est proche en termes de direction (voir figure 4.2) ainsi la modification se fera automatiquement entre des modes d'une même tranche, en plus nous proposons dans certains cas s'insérer 2 bits par bloc pour augmenter la capacité d'insertion comme expliquer dans la prochaine section.

4.7 L'approche proposée

4.7.1 Présentation générale

L'insertion de données secrètes est effectuée dans les images I, dans les modes de prédiction Intra 4x4.

En observant les neuf modes de prédiction (figure 4.2), il serait possible de remarquer qu'ils ont des directions proches qui utilisent les mêmes échantillons voisins dans la prédiction. Dans ce cas, il serait possible d'éviter les problèmes de disponibilité du mode si l'insertion était basée sur la modification des modes de directions proches.

En se basant sur cette idée, nous pouvons définir des groupes de modes comme suit:

Groupe 1: Modes 1 et 8.

Groupe 2: Modes 3 et 7.

Groupe 3: Modes 2, 4, 5 et 6.

Groupe 4: Modes 0 et 2.

Le mode DC est rajouté au troisième groupe afin de renforcer la capacité d'insertion. Le groupe 4 est créé afin d'exploiter le mode vertical notamment pour les séquences dans lesquelles ce mode serait dominant.

Cependant, pour pallier le problème de non disponibilité des échantillons en cas de remplacement du mode 2 par l'un des modes 4, 5 ou 6, nous proposons de créer un vide dans le mode 2 comme suit: Le mode 2 est changé par le mode le plus probable (MPM) dans le cas où MPM est différent de 2, autrement il est changé par le mode 0 ou 1 selon la disponibilité.

La figure 4.3 montre l'histogramme des modes de prédiction Intra 4x4 de la première image de la séquence Bridge-Close avant et après la création du vide.

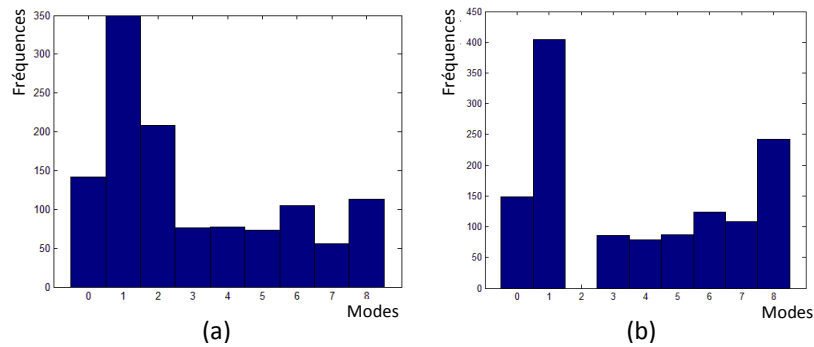


Figure 4.3 : Histogramme des modes de prédiction Intra 4x4 d'une image I de la séquence Bridge-Close: (a) Etat initial, (b) Après création du vide dans le mode DC.

La modification des modes est validée par l'optimisation du taux de distorsion comme présenté dans la sous-section suivante, autrement l'encodeur utilise le mode intra 16x16 comme expliqué dans la section 4.8 (voir figure 4.5).

4.7.2 Processus d'insertion

Le processus d'insertion est appliqué aux modes de prédiction comme suit: Pour le groupe 1, le mode horizontal (1) est inchangé quand le bit inséré est égal à 1 et il est égal au mode 8 lorsque le bit est 0. La même chose pour le mode 8, qui reste inchangé lorsque le bit inséré est 0, et il est modifié à 1 si le bit est égal à 0. Pour chaque cas (MM=1 ou MM=8), l'insertion n'est pas exécutée si le MPM est égal à 1 ou 8 en vue de limiter l'augmentation du débit et d'assurer une détection fiable. Par exemple, si le mode 1 est modifié à 8, et le MPM est égal à 8, lors de l'extraction, le mode 8 serait ignoré car il correspondrait au MPM et la détection serait erronée.

Le même principe est appliqué au groupe 2, comme mentionné sur le tableau 4.2.

Tableau 4.2. Configuration des modes pour l'insertion de la marque pour les groupes 1 et 2.

Bits de la marque	Modification des modes	
	3 et 7	1 et 8
0	7	8
1	3	1

Pour les modes 4, 5 et 6, deux bits sont insérés tel que représenté sur le tableau 4.3. Par exemple, si le meilleur-mode est égal à 4 et le MPM est différent de 4, 5 ou 6, le mode 4 devient 2 si deux bits successifs de la marque sont "00", et le mode devient égal à 6 si les bits successifs de la marque sont "11".

Enfin, si le meilleur-mode est égal au mode 0 et le MPM est différent de 0, le meilleur mode est modifié en mode DC uniquement si les bits de la marque W_i et $W_i + 1$ sont égaux à 0, sinon, le meilleur mode n'est pas changé. Cela se justifie par la disponibilité du mode DC tel que mentionné précédemment, et aussi par le fait que le mode vertical ne peut pas être changé par les autres modes, sans vérification de la disponibilité des modes et sans risquer une augmentation importante du débit.

Tableau 4.3. Configuration des modes pour l'insertion de la marque pour les groupes 3 et 4.

Bits de la marque	Modification des modes	
	4, 5 et 6	0
00	2	2
01	4	Non utilisé
10	5	
11	6	

Un exemple est donné ci-dessous, relatif au code utilisé pour l'insertion de deux bits en changeant le mode de la diagonale bas-droite (mode 4). En utilisant le code source JM8.6¹¹, le code est intégré dans le fichier de `rdopt.c`, dans la fonction de sélection du mode (`Mode_Decision_for_4x4IntraBlocks`) et à l'intérieur de la boucle de contrôle du taux de distorsion.

```

if (img->type == I_SLICE) {
if (ipmode==4 && MPM != 4 && MPM != 5 && MPM != 6 )
{
if ( ipmode==4 && (W[i]==0 && W[i+1]==0 ))
{
ipmode=2;
GetPredictionError ( ); Rdcost(); IncrementWatermark( );
}
else if (ipmode==4 && (W[i]==0 && W[i+1]==1 ))
{
IncrementWatermark( ); }
else if (ipmode==4 && (W[i]==1 && W[i+1]==0 ))
{
ipmode=5;
GetPredictionError ( ); Rdcost(); IncrementWatermark( );
}
else if ( ipmode==4 && (W[i]==1 && W[i+1]==1))
{
ipmode=6;
GetPredictionError( ); Rdcost(); IncrementWatermark( );
}
}
}
}

```

Le code est appliqué uniquement si les images sont de type I (`I_SLICE`).

`ipmode` est le mode en cours qui est le meilleur mode, `W` est la marque à insérer.

¹¹ Iphome.hhi.de/suehring/tml/

IncrementWatermark() est une fonction utilisée pour incrémenter la marque par deux bits. GetPredictionError() renvoie aux deux opérations qui existent déjà dans la fonction (*Mode_Decision_for_4x4IntraBlocks*) et qui sont utilisées pour calculer la prédiction de bloc et la différence entre l'original et le bloc prédit. Rdcost () représente la fonction qui calcule le coût du taux de distorsion (*Rate Distorsion Cost*).

4.7.3 Processus d'extraction

À l'étape du décodage et après le décodage entropique, le processus d'extraction peut se faire facilement après le décodage entropique, puisque nous pouvons en déduire les valeurs des bits insérés selon les modes tels que présentés dans les tableaux 4.2 et 4.3. Bien sûr, à cette étape nous maintenons la condition que pour un mode donné, le mode le plus probable ne doit pas être égal à ce mode ou un autre du même groupe, excepté pour le mode DC, qui subit un traitement spécial, comme expliqué dans la section 4.7.1.

Un exemple de code d'extraction de bits est donné ci-dessous:

```

if (mode==1 && MPM!= 1 && MPM != 8 ) {W[i]=1; i=i+1; }
if (mode==8 && MPM!= 1 && MPM != 8 ) {W[i]=0; i=i+1; }
if (mode==3 && MPM!= 3 && MPM != 7 ) {W[i]=1; i=i+1; }
if (mode==7 && MPM!= 3 && MPM != 7 ) {W[i]=0; i=i+1; }

if (mode==2 ) {W[i]=0;W[i+1]=0; i=i+2; }
if (mode==4 && MPM!= 4 && MPM != 5 && MPM!= 6 ) {W[i]=0;W[i+1]=1; i=i+2; }
if (mode==5 && MPM!= 4 && MPM != 5 && MPM!= 6 ) {W[i]=1;W[i+1]=0; i=i+2; }
if (mode==6 && MPM!= 4 && MPM != 5 && MPM!= 6 ) {W[i]=1;W[i+1]=1; i=i+2; }

```

4.7.4 La sécurité et la robustesse du tatouage

Dans le processus d'insertion décrit précédemment, la position d'insertion est connue et les bits cachés peuvent être facilement extraits. Ainsi, la sécurité de l'information secrète pourrait être assurée par des outils de chiffrement qui peuvent être effectuées lors d'une étape de prétraitement. Par ailleurs, il est possible de changer la configuration utilisée pour l'insertion de la marque (tableaux 4.2 et 4.3) et la considérer comme une information secrète, mais elle doit être transmise comme une information additionnelle. Nous pouvons également utiliser une sélection aléatoire de blocs pour renforcer l'aspect de sécurité, mais dans ce cas la capacité d'insertion serait réduite.

Pour la robustesse de l'approche, toute attaque du flux vidéo causerait des erreurs à l'étape de détection et affecterait le contenu de la vidéo, ce qui constitue une preuve que la vidéo a été attaquée. Cependant, les bits insérés peuvent être perdus dans le cas où nous effectuons un décodage puis un ré-encodage de la séquence vidéo, pour cette raison l'approche proposée est

plus adaptée pour des applications telles que la communication secrète ou l'authentification de la vidéo compressée plutôt que la protection du copyright. L'amélioration de la robustesse et de la sécurité pourrait être obtenue, au détriment de la capacité d'insertion, en modulant le message à insérer par une séquence binaire pseudo aléatoire [22].

4.8 Résultats et analyses

La méthode décrite précédemment a été testée sur huit séquences vidéo de format QCIF: Bridge-Close, Carphone, Coast-Guard, Container, Foreman, Grandma, News et Silent. Pour chaque séquence, 200 images ont été codées à une cadence de 30 images / seconde. Le profil utilisé est le principal et la structure du groupe d'images (GOP) est "IBPBPBPBPB".

Pour les tests, nous avons utilisé une séquence binaire pseudo aléatoire comme marque. L'insertion a été effectuée dans les images de type I en intégrant l'algorithme dans le code source du codec H.264/AVC. Les résultats sont présentés dans le tableau 4.4.

Tableau 4.4. Les résultats obtenus pour huit séquences vidéo.

Séquences vidéo	Capacité (bits)	Débit (%)	Red. du PSNR (dB)
Foreman	16291	5.19	0.06
Grandma	11412	5.11	0.06
Container	8331	5.50	0.05
Coast-Guard	9974	2.18	0.05
Carphone	13436	4.03	0.05
Silent	17716	5.72	0.05
News	11645	4.45	0.05
Bridge-Close	12007	3.26	0.04



Figure 4.4: Image I de la séquence Container : Macroblocs codés en Intra 4x4 et en Intra 16x16.

La capacité d'insertion dépend du nombre de macroblocs codés en intra 4x4 et également de l'existence de modes qui transportent 2 bits (tels que 4, 5 et 6). Les résultats montrent que la plus faible capacité d'insertion est 8331 bits, obtenue pour la séquence Container qui se caractérise par peu de détails. En effet, la séquence Container contient un nombre relativement important de

macroblochs codés en intra 16x16 (voir figure 4.4) et qui ne sont pas utilisés pour l'insertion des données.

La plus grande capacité d'insertion est de 17716 bits obtenu pour la séquence Silent qui contient, au contraire, une quantité de détails importante. La capacité d'insertion concerne les 20 I images contenues dans la séquence.

Nous avons implémenté la méthode proposée dans [8] dans les mêmes conditions. Les résultats montrent que la capacité d'insertion que nous avons obtenue est plus importante (figure 4.5) pour une pareille augmentation de débit.

Nous avons aussi comparé notre approche aux méthodes proposées dans [6][7], la capacité d'insertion est plus élevée comme le montre la figure 4.5. Toutefois, selon les résultats, cette méthode ne serait pas recommandée pour les séquences avec peu de détails comme Grandma et Container, qui ont relativement un faible rapport capacité/débit par rapport aux séquences contenant plus de détails tels que Silent ou Foreman.

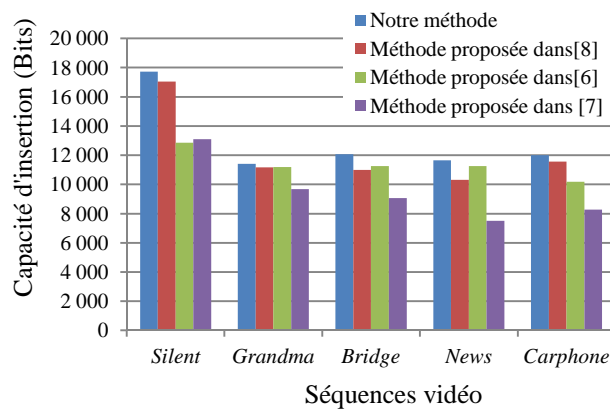


Figure 4.5 : Capacité d'insertion pour cinq séquences tatouées avec notre méthode et avec les méthodes proposées par Yang, Hu et Xu.

Comme le montre le tableau 4.3, le débit est relativement élevé. Il varie entre 2.18 et 5.72% enregistré pour la séquence Silent caractérisée par une capacité d'insertion relativement élevée. Ainsi, cette approche peut ne pas convenir pour des applications exigeantes en termes de débit. A titre indicatif, la méthode proposée dans [7] présente des résultats liés à l'augmentation du débit, qui varie entre 2.3 et 3.3% pour les séquences mentionnées dans la figure 4.4, alors que les résultats sont moins de 2% pour la méthode proposée dans [6], pour les mêmes séquences.

Cette augmentation de débit était prévisible parce que la méthode proposée n'est pas basée sur le changement du mode optimal par le mode sous-optimal comme expliqué précédemment. En effet, le contrôle du débit est uniquement basé sur l'insertion des données secrètes lorsque le meilleur mode est différent du mode le plus probable. Cependant, notre objectif était d'éviter d'utiliser toutes les configurations (les mises en correspondance) qui exigeraient une

connaissance préalable des caractéristiques des images, ou l'encodage préalable de la vidéo, en proposant une méthode qui permet d'obtenir une insertion simple, facile à mettre en œuvre et approprié pour une application en temps réel comme mentionné ultérieurement. Au même temps, une capacité d'insertion relativement élevée est assurée en particulier pour les séquences hautement texturées.

La figure 4.6 montre que l'insertion des données, qui est contrôlée par le coût du taux de distorsion, peut augmenter le nombre de macroblocs codés en intra 16x16, et donc, a un effet sur la capacité d'insertion et l'augmentation du débit.

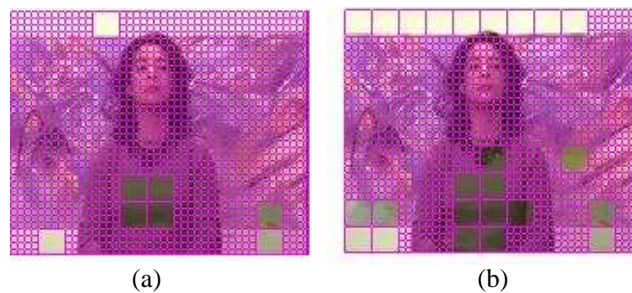


Figure 4.6 : Augmentation du nombre de macroblocs codés en intra 16x16 dans la première image I de la séquence Silent: (a) la séquence non marquée, (b) la séquence marquée.

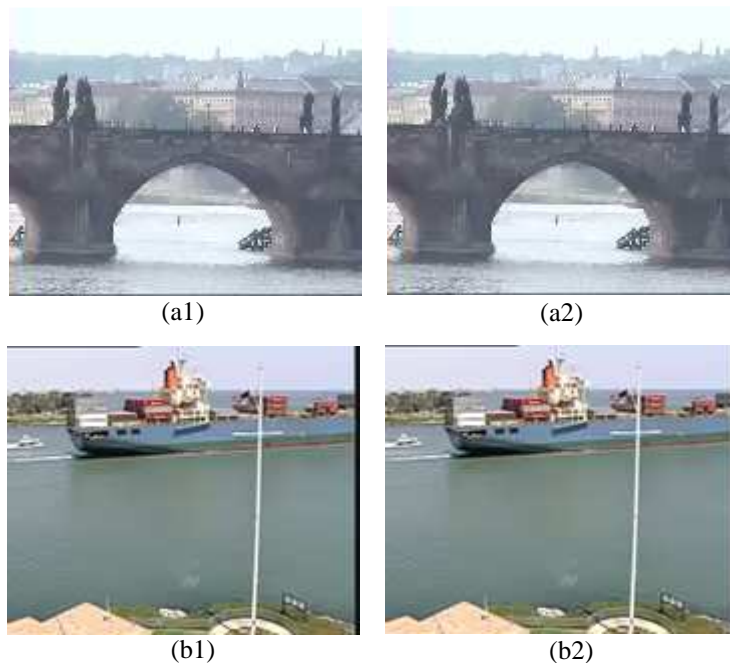


Figure 4.7 : (a1) et (a2) représentent les images I non marquées des séquences Bridge et Container, et (b1) et (b2) sont les images marquées.

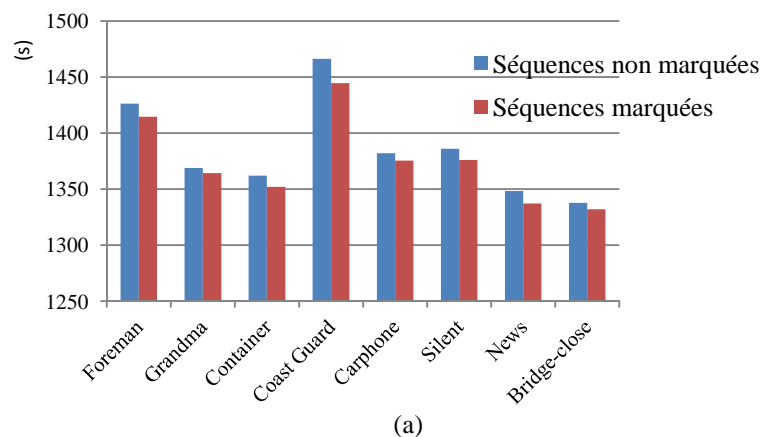
L'évaluation objective de la vidéo tatouée mesurée par le PSNR a montré que la qualité vidéo est maintenue car la diminution des valeurs de PSNR, en comparant avec les séquences non marquées, sont de l'ordre de centièmes (par exemple, pour la séquence Container, le PSNR a diminué de 0.03 dB). La qualité subjective est parfaitement préservée comme le montre la

figure 4.7 qui représente les images I de Container et de Bridge-Close avant et après l'insertion des données.

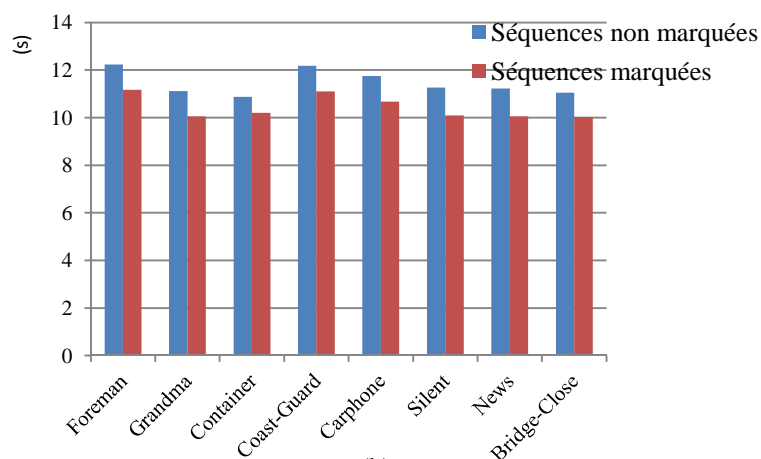
L'étape de détection est fiable pour toutes les séquences de test sans appliquer aucune attaque. L'extraction des données peut se faire facilement sans un décodage complet de la séquence, après l'étape de décodage entropique comme expliqué à la section 4.7.3.

Tableau 4.5. Encodage des premières images de la séquence Silent non marquée et marquée.

Image	Séquence Silent non marquée		Séquence Silent marquée	
	Temps (ms)	MET (ms)	Temps (ms)	MET (ms)
0000(I)	702	0	726	0
0002(P)	1138	979	1170	250
0001(B)	2496	1007	2496	1036
0004(P)	1607	624	1544	421
0003(B)	3416	1009	3369	1579
0006(P)	2060	781	2074	780
0005(B)	4337	2060	4258	2224
0008(P)	2558	1073	2558	1210
0007(B)	5179	2609	5148	2762
0010(I)	687	0	718	0
0009(B)	6131	3226	6053	3181



(a)



(b)

Figure 4.8 : (a) et (b) représentent le temps d'encodage et de décodage des séquences de tests marquées et non marquées.

Les résultats ont montré que le temps d'encodage des images I augmente lors de l'insertion des données secrètes, alors que le temps d'encodage de P ou B peut être maintenu ou réduit comme indiqué dans le tableau 4.5, dans lequel le temps d'encodage peut être comparé entre la séquence Silent non marquée et marquée. La désignation «Temps» indique le temps d'encodage total et "MET" indique le temps d'estimation de mouvement total pour les images. La figure 4.8.a présente les résultats obtenus pour les temps d'encodage. Les tests ont été effectués avec un PC Pentium 4.3 GHz, 1 Go de Ram utilisant le système d'exploitation Windows XP.

Réalisée avec les mêmes performances de l'ordinateur, les résultats ont montré que le temps de décodage (extraction de données) est maintenu ou réduit, comme illustré sur la figure 4.7(b).

4.9 Conclusion

Dans ce chapitre, une nouvelle approche d'insertion des données secrètes en exploitant des modes de prédiction intra 4x4 a été présentée. Les résultats ont montré qu'une capacité d'insertion élevée peut être atteinte, notamment pour les séquences hautement texturées, tandis que les temps d'encodage et de décodage ne sont pratiquement pas perturbés.

Par conséquent, l'utilisation des modes de prédiction intra est en effet motivante pour toutes les possibilités qu'elle offre en termes de préservation de la qualité vidéo, l'amélioration de la capacité d'insertion et le maintien des temps d'encodage et de décodage. Cependant, des améliorations devraient être apportées à cette approche afin de contrôler l'augmentation du débit.

L'utilisation de la prédiction intra pour insérer des données secrètes dans le codec H.264/AVC a donné de meilleurs résultats par rapport aux méthodes basées sur les coefficients DCT [69]. Toutefois, d'importantes améliorations pourraient être apportées à ses dernières si le problème de la propagation d'erreur était traité.

Dans le prochain chapitre nous nous intéressons au problème de la propagation de l'erreur qui apparaît au niveau du décodeur lors de l'insertion de la marque dans les coefficients DCT. L'objectif est de pouvoir augmenter la capacité d'insertion tout en améliorant la qualité visuelle.

5.1 Introduction

Bien que les résultats des techniques de tatouage basées sur les modes de prédiction soient satisfaisants, les méthodes basées sur les coefficients quantifiés restent plus intéressantes pour la communauté des chercheurs étant donné qu'elles donnent plus de choix dans la manipulation des coefficients et par conséquent différentes caractéristiques peuvent être obtenues, telles que la réversibilité, la robustesse et la sécurité du tatouage [70].

Toutefois, les méthodes de tatouage basées sur les coefficients DCT quantifiés sont plus concernées par le problème de la propagation de l'erreur qui apparaît lors du décodage du flux vidéo. Ce problème freine leur efficacité et cause la propagation de la dégradation des blocs encodés à toute la séquence vidéo à travers la prédiction intra et inter-images.

Aujourd'hui, il existe peu de travaux qui traitent du problème de la propagation d'erreur pour le codec H.264/AVC. Dans [71], les auteurs ont proposé un schéma qui cible la rapidité et la robustesse. La marque est insérée en modifiant les coefficients DC quantifiés dans les blocs de résidus luma. Un masquage textural est utilisé pour choisir la force adaptée au marquage pour chaque bloc, l'objectif est d'assurer la robustesse et de maintenir la qualité perceptuelle de la vidéo. Pour éliminer l'effet de la propagation de l'erreur, un signal de compensation est calculé pour chaque bloc avant l'insertion de la marque. Toutefois, dans ce schéma, le message secret est modulé par une séquence bipolaire qui réduit la quantité de données secrètes insérées. De plus, la marque est nécessaire pour effectuer la détection et elle est insérée directement dans le flux binaire (après la phase de compression), impliquant ainsi des opérations supplémentaires (décodage et ré-encodage entropique), ce qui pourrait être un obstacle pour une diffusion de la vidéo en temps réel.

Dans [9], les auteurs ont proposé un schéma de marquage qui répond au problème de la propagation de l'erreur en intra. Les données secrètes sont insérées dans les images de type I, dans les coefficients DCT 4x4 de luminance quantifiées. Des paires de coefficients sont définies dans les blocs de luminance 4x4 pour prévenir la propagation de la distorsion. L'insertion des données dans les deux coefficients permet d'amener l'erreur introduite au milieu de la matrice du bloc.

Dans cette méthode, tel qu'expliqué dans la section 5.3, le flux H.264 est d'abord décodé, pour chaque bloc candidat au marquage. La paire de coefficients est sélectionnée selon le mode de prédiction des blocs voisins. Les blocs sont ensuite ré-encodés. Le principal inconvénient de cette méthode réside dans le fait qu'elle agit sur le flux vidéo et nécessite une connaissance

préalable des modes de prédiction des blocs voisins, ce qui constitue également une contrainte pour la diffusion en temps réel.

Dans [72], les auteurs ont proposé de supprimer la propagation d'erreur en calculant la différence entre les échantillons tatoués et originaux et faire une compensation pour tous les coefficients DCT entiers. Cette méthode a été améliorée dans [73], en proposant trois algorithmes qui compensent seuls les blocs marqués selon la position du coefficient affecté. L'objectif principal est de réduire la charge de calcul de cette approche. Cependant, cet algorithme est « détectable », c'est à dire qu'il nécessite la connaissance de la signature pour la détection de la donnée secrète, ce qui limiterait le nombre d'applications.

Dans ce chapitre, une nouvelle solution est étudiée pour compenser la propagation des erreurs. La méthode est basée sur l'estimation de l'erreur introduite dans le bloc marqué ou bien celle propagée dans les blocs voisins. Deux schémas sont présentés [70][74]: Dans le premier algorithme, la méthode proposée dans [9] est adaptée pour être utilisée au cours du processus d'encodage. La distorsion causée par chaque paire de coefficients est calculée de manière à donner une priorité aux paires de tatouage qui introduisent une erreur minimale. Dans le deuxième schéma, l'erreur estimée dans les blocs voisins est réduite à partir des résidus au cours du processus de codage. Dans les deux schémas, le processus d'insertion adopté joue un rôle crucial dans le maintien du débit de la vidéo.

5.2 La propagation de l'erreur

5.2.1 Les processus de codage et de décodage du codec H.264/AVC

Rappelons brièvement le processus de codage vidéo (voir chapitre 1): les principales étapes d'encodage sont la prédiction, la transformation et la quantification, et le codage entropique. Dans la norme H.264, la première image d'une scène est codée en intra: un macrobloc (16x16 échantillons) d'intérêt est prédit à partir d'un macrobloc déjà codé et situé à gauche ou au-dessus. La différence entre le macrobloc (ou bloc) original et celui prédit R représente des données résiduelles qui subissent une transformation, une quantification et un codage entropique. Outre la prédiction intra, une prédiction inter-image est exploitée entre des images adjacentes pour réduire la redondance temporelle.

À l'étape de décodage, un macrobloc est déduit à partir des données résiduelles et de la prédiction. Les coefficients quantifiés L (appelés aussi niveaux) sont récupérés à ce stade après le décodage entropique. Après la mise à l'échelle et l'opération de transformation inverse, les données résiduelles R' sont déduites. À la figure 5.1 (a) et (b) est représenté un schéma général des opérations de codage et de décodage.

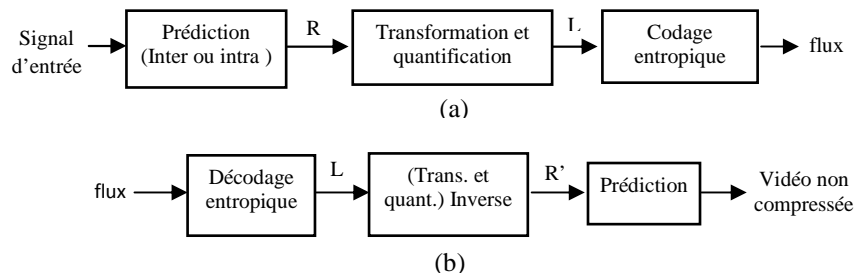


Figure 5.1 : (a) et (b) représentent un schéma général des processus D'encodage et de décodage respectivement.

5.2.2 Description de la propagation de la distorsion

Au niveau de l'encodage, le bloc de luminance A peut être prédit selon neuf modes de prédiction en utilisant les blocs voisins précédemment codés, et le meilleur mode est déduit à partir de la technique de l'optimisation du taux de distorsion [3]. Le troisième mode est montré comme exemple à la figure 5.2(a). Pour ce mode, les seize échantillons prédits ($Prediction[i][j]$, $i=1,\dots,4$, $j=1,\dots,4$) du bloc A sont déduits des échantillons de a à h comme suit [7]:

$$Prediction[0][0] = (a + c + 2*b + 2) / 4,$$

$$Prediction [0][1] = Prediction [1][0] = (b + d + 2*c + 2) / 4,$$

$$Prediction [0][2] = Prediction [1][1] = Prediction [2][0] = (c + e + 2*d + 2) / 4,$$

$$Prediction [0][3] = Prediction [1][2] = Prediction [2][1] = Prediction [3][0] \\ = (d + f + 2*e + 2) / 4,$$

$$Prediction [1][3] = Prediction [2][2] = Prediction [3][1] = (e + g + 2*f + 2) / 4,$$

$$Prediction [2][3] = Prediction [3][2] = (f + h + 2*g + 2) / 4,$$

$$Prediction [3][3] = (g + 3*h + 2) / 4,$$

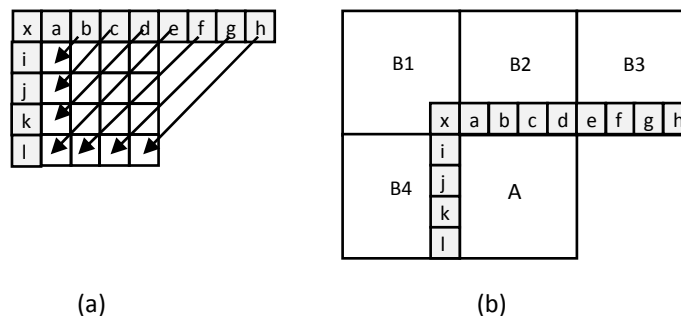


Figure 5.2 : (a) Mode de Prédiction n°3, (b) Blocs et échantillons de prédiction.

À l'étape du décodage, les blocs sont reconstruits à partir des blocs prédits et des résidus de R_A' . Si un bit secret est ajouté à l'un des blocs B1, B2, B3 ou B4 (figure. 5.2 (b)) en modifiant l'un des seize coefficients, les données résiduelles R_{B_i}' ($i = 1, \dots, 4$), et en particulier les échantillons x, a, ... l seraient modifiés, et la prédiction de A serait également changée. De cette façon, une

erreur introduite dans un bloc peut se propager à toute l'image à travers la prédiction intra, et à toute la séquence vidéo à travers la prédiction inter (voir section 1.3.3).

5.2.3 L'estimation de l'erreur

Au niveau de l'encodeur, la DCT entière d'un bloc de résidus X est donnée par [3]:

$$Y = C_f X C_f^T \otimes E_f \quad (5.1)$$

$$W_{ij} = C_f X C_f^T \quad (5.2)$$

Le symbole \otimes représente la multiplication scalaire entre W_{ij} et la matrice des facteurs d'échelle E_f .

$$C_f = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & -1 & 2 \end{bmatrix}$$

Les coefficients quantifiés sont obtenus par (5.3), qui effectue la quantification et la mise à l'échelle en une seule opération:

$$Z_{ij} = \text{round} \left(\frac{W_{ij} PF}{Qstep} \right) \quad (5.3)$$

$Qstep$ est le pas de quantification et PF est égal à a^2 , $ab/2$ or $b^2/4$ ($a=1/2$, $b= \sqrt{2}/5$) selon la position de (i,j) .

$$PF = \begin{bmatrix} a^2 & ab/2 & a^2 & ab/2 \\ ab/2 & b^2/2 & ab/2 & b^2/2 \\ a^2 & ab/2 & a^2 & ab/2 \\ ab/2 & b^2/2 & ab/2 & b^2/2 \end{bmatrix}$$

Au niveau du décodeur, les coefficients résiduels X'_{ij} sont obtenus par (5.4) et (5.5) qui représentent l'étape de remise à l'échelle et la DCT entière inverse respectivement. Le bloc de pixels est reconstruit en ajoutant le bloc de résidus au bloc prédit.

$$W'_{ij} = Z_{ij} Qstep . PF' . 64 \quad (5.4)$$

$$X'_{ij} = \text{round} [(C_i^T W'_{ij} C_i) / 64] \quad (5.5)$$

$$C_i = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1/2 & -1/2 & -1 \\ 1 & -1 & -1 & 1 \\ 1/2 & -1 & 1 & -1/2 \end{bmatrix}$$

PF' est égal à a^2 , ab or b^2 selon la position de (i,j) .

Dans les techniques de marquage basées sur les coefficients DCT, les données secrètes sont insérées en ajoutant une matrice d'erreur Δ (composées des coefficients Δ_{ij}) aux coefficients quantifiés. Les nouveaux coefficients mis à l'échelle deviennent :

$$W_{ij}'' = (Z_{ij} + \Delta_{ij}) Qstep PF' 64 \quad (5.6)$$

Les coefficients résiduels sont obtenus par:

$$X_{ij}'' = \text{round} [((C_i^T (Z_{ij} Qstep PF' 64) C_i) + (C_i^T (\Delta_{ij} \cdot Qstep PF' 64) C_i)) / 64] \quad (5.7)$$

5.3 Limitation de la propagation d'erreur basée sur le marquage de paires de coefficients

5.3.1 Elimination de la propagation d'erreur

Ma et al.[9] ont proposé une méthode basée sur le marquage de paires de coefficients. Les auteurs ont étudié l'étape de remise à l'échelle, et de quantification et transformation inverses pour identifier des paires de coefficients qui apportent l'erreur au milieu de la matrice du bloc de coefficient. Ces paires de coefficients sont rassemblées en deux ensemble : l'ensemble horizontal et l'ensemble vertical comme suit :

L'ensemble horizontal (HS): $\{(Z_{01}, Z_{21}), (Z_{21}, Z_{01}), (Z_{02}, Z_{22}), (Z_{22}, Z_{02}), (Z_{03}, Z_{23}), (Z_{23}, Z_{03})\}$

L'ensemble vertical (VS): $\{(Z_{10}, Z_{12}), (Z_{12}, Z_{10}), (Z_{20}, Z_{22}), (Z_{22}, Z_{20}), (Z_{30}, Z_{32}), (Z_{32}, Z_{30})\}$

Toutes les paires de coefficients de HS et VS accumule la distorsion au centre de la matrice en ligne ou en colonne. Un exemple est donné dans le tableau 5.1, où Δ' est la matrice d'erreur composée de Δ'_{ij} , déduits de (5.7):

$$\Delta'_{ij} = (C_i^T (\Delta_{ij} \cdot Qstep PF' 64) C_i) / 64 \quad (5.8)$$

En se basant sur cette idée, un bit inséré dans une paire de coefficient de HS pourrait éviter la propagation de l'erreur dans les blocs voisins B, C et D s'ils utilisaient les échantillons m, n, o et p (figure 5.3) dans la prédiction intra. De même, un bit inséré dans une paire de coefficient VS pourrait éviter la propagation de l'erreur aux blocs A et D s'ils utilisaient les échantillons d, h, l et p dans leur prédiction.

Pour insérer un bit, les auteurs proposent la sélection de paires de coefficients selon des conditions qui éviteraient la propagation de l'erreur. Une connaissance préalable des modes de prédiction de A, B, C et D (figure 5.3) est nécessaire pour appliquer cette méthode. Pour cela, elle ne peut être appliquée que sur le flux H.264 et non pas durant le processus d'encodage. En plus, si les deux coefficients d'une paire sont nuls, l'insertion n'est pas effectuée ce qui réduit la capacité d'insertion, par ailleurs, l'insertion peut être effectuée si l'un des coefficients de la paire

est nul, ce qui augmenterait considérablement le débit. Pour pallier ces inconvénients, nous proposons une adaptation de cette méthode comme présenté dans la prochaine section.

Tableau 5.1. Exemple de matrice d'erreur dans le cas d'insertion d'un bit dans un coefficient ou une paire de coefficients.

Δ (Données insérée dans un bloc)	Δ' (Matrice d'erreur au niveau du décodeur)
$\begin{bmatrix} 0 & 0 & t & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} a^2 t Q_{step} & -a^2 t Q_{step} & -a^2 t Q_{step} & a^2 t Q_{step} \\ a^2 t Q_{step} & -a^2 t Q_{step} & -a^2 t Q_{step} & a^2 t Q_{step} \\ a^2 t Q_{step} & -a^2 t Q_{step} & -a^2 t Q_{step} & a^2 t Q_{step} \\ a^2 t Q_{step} & -a^2 t Q_{step} & -a^2 t Q_{step} & a^2 t Q_{step} \end{bmatrix}$
$\begin{bmatrix} 0 & 0 & t & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -t & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 2a^2 t Q_{step} & -2a^2 t Q_{step} & -2a^2 t Q_{step} & 2a^2 t Q_{step} \\ 2a^2 t Q_{step} & -2a^2 t Q_{step} & -2a^2 t Q_{step} & 2a^2 t Q_{step} \\ 0 & 0 & 0 & 0 \end{bmatrix}$
$\begin{bmatrix} 0 & 0 & 0 & 0 \\ t & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} abtQ_{step} & abtQ_{step} & abtQ_{step} & abtQ_{step} \\ abtQ_{step}/2 & abtQ_{step}/2 & abtQ_{step}/2 & abtQ_{step}/2 \\ -abtQ_{step}/2 & -abtQ_{step}/2 & -abtQ_{step}/2 & -abtQ_{step}/2 \\ -abtQ_{step} & -abtQ_{step} & -abtQ_{step} & -abtQ_{step} \end{bmatrix}$
$\begin{bmatrix} 0 & 0 & 0 & 0 \\ t & 0 & -t & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 2abt Q_{step} & 2abt Q_{step} & 0 \\ 0 & abt Q_{step} & abt Q_{step} & 0 \\ 0 & -abt Q_{step} & -abt Q_{step} & 0 \\ 0 & -2abt Q_{step} & -2abt Q_{step} & 0 \end{bmatrix}$

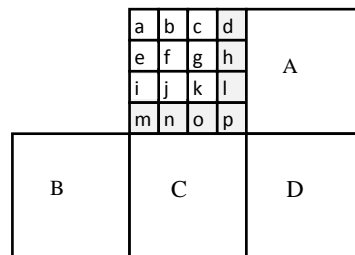


Figure 5.3 : Blocs voisins se positionnant à droite et en bas du bloc en cours

5.3.2 Approche proposée: Premier schéma

Dans cette approche, l'insertion des données secrètes se fait dans les images de type Intra, Durant le processus d'encodage du H.264/AVC.

Les blocs candidats au marquage sont les blocs DCT 4x4 avec relativement un grand nombre de résidus. La sécurité des données secrètes est basée sur des techniques de cryptage.

L'insertion est appliquée dans les paires de coefficients dans le but de ramener l'erreur vers le milieu de la matrice tel que mentionné précédemment. Six paires de coefficients peuvent être utilisées : $(Z_{01}, Z_{21}), (Z_{02}, Z_{22}), (Z_{03}, Z_{23}), (Z_{10}, Z_{12}), (Z_{20}, Z_{22}), (Z_{30}, Z_{32})$, correspondant aux

positions suivantes dans le balayage en zigzag : (P_1, P_8) , (P_5, P_{11}) , (P_6, P_{13}) , (P_2, P_7) , (P_3, P_{11}) , (P_9, P_{14}) comme illustré dans la figure 5.4 (a).

De la formule (5.7), il est possible d'estimer à l'avance l'erreur introduite dans chaque bloc. En effet si l'insertion des données secrètes est basée sur un ajout ou une soustraction de 1 aux coefficients quantifiés, il est possible de calculer Δ'_{ij} tel qu'indiquer dans la section 5.3.1.

La figure 5.4 (de (b) à (g)) présente les résultats obtenus en insérant un bit aux paires de coefficients (en ajoutant 1 à un coefficient et en soustrayant 1 au second) des positions : (P_1, P_8) , (P_2, P_7) , (P_3, P_{11}) , (P_5, P_{11}) , (P_6, P_{13}) et (P_9, P_{14}) . Seulement les échantillons qui jouent un rôle dans la propagation de l'erreur ont été mentionnés.

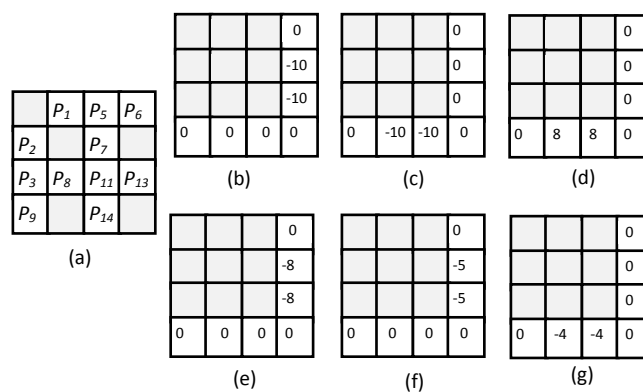


Figure 5.4 : (a) Représentation des positions des paires de coefficients, (b), (c), (d), (e), (f) et (g) L'estimation de l'erreur introduite dans les échantillons de bloc en insérant un bit dans (P_1, P_8) , (P_2, P_7) , (P_3, P_{11}) , (P_5, P_{11}) , (P_6, P_{13}) et (P_9, P_{14}) respectivement.

En se basant sur cette estimation, il est possible de donner des priorités de marquage aux paires qui introduisent une erreur de propagation minimale aux blocs voisins. Les paires de coefficients sont triés comme suit : (P_9, P_{14}) introduit l'erreur minimale, en suite (P_6, P_{13}) , et $((P_5, P_{11})$ ou $(P_3, P_{11}))$, et finalement $((P_2, P_7)$ ou $(P_1, P_8))$.

Pour le processus d'insertion, un bloc candidat est d'abord sélectionné selon le nombre de N des résidus (pour les expérimentations, trois cas sont pris en considération: $N > 5$, $N > 7$ et $N > 9$).

Pour chaque bloc, les paires de coefficients non nuls sont identifiées et un bit est inséré dans la paire adéquate selon la priorité et la disponibilité.

Il est nécessaire de choisir les valeurs de coefficients qui n'affectent pas la qualité de la vidéo et son débit tels que les coefficients 0 et 1 qui jouent un rôle important dans l'étape du codage entropique. La modification du coefficient 0 augmente considérablement le débit, la modification du 1 (à 0) affecte considérablement la qualité de la vidéo, même si cela a l'avantage de réduire le débit. Pour l'insertion d'un bit W nous appliquons la modulation représentée sur le tableau 5.2, où (C_1, C_2) est la paire de coefficient, $C_1 \neq 0$ et $C_2 \neq 0$.

Dans le cas où chacun des coefficients C_1 et C_2 est égal à 1 ou -1, cette paire n'est pas utilisée pour l'insertion afin de ne pas transformer un coefficient en 0. Si la paire (P_5, P_{11}) n'est pas utilisée pour cette raison, la paire (P_3, P_{11}) n'est également pas utilisée car le résultat de détection pourrait être erroné à cause de la position commune P_{11} .

Tableau 5.2. Processus d'insertion

Conditions		Instructions	
(C ₁ mod 2 = 0 and W=1) ou (C ₁ mod 2 = 1 and W=0)	C ₁ > 0	C ₂ ≠ 1	C ₁ = C ₁ + 1, C ₂ = C ₂ - 1
		C ₂ = 1, C ₁ ≠ 1	C ₁ = C ₁ - 1, C ₂ = C ₂ + 1
	C ₁ < 0	C ₂ ≠ -1	C ₁ = C ₁ - 1, C ₂ = C ₂ + 1
		C ₂ = -1, C ₁ ≠ -1	C ₁ = C ₁ + 1, C ₂ = C ₂ - 1
(C ₁ mod 2 = 0 and W=0) ou (C ₁ mod 2 = 1 and W=1)	C ₁ et C ₂ ne sont pas modifiés.		

L'étape d'extraction n'a pas besoin d'un décodage complet du flux vidéo. Après le décodage entropique, les blocs de coefficients DCT quantifiés avec un nombre N de coefficients non nuls sont sélectionnés et la disponibilité des coefficients (C_1, C_2) de la position (P_9, P_{14}) est vérifiée. Si au moins $\text{abs}(C_1) \neq 1$ ou bien $\text{abs}(C_2) \neq 1$, le bit W est extrait du bloc comme suit:

si $C_1 \text{ mod } 2 = 1$, $W = 1$, sinon, $W = 0$.
 abs est la valeur absolue et mod est l'opération de modulo.

Si une paire n'est pas disponible ($C_1=0$ ou $C_2=0$) ou ne remplit pas les conditions de marquage : ($\text{abs}(C_1) = 1$ et $\text{abs}(C_2) = 1$), la même vérification est effectuée pour la paire suivante jusqu'à extraire toutes les données cachées.

5.4. Limitation de la propagation d'erreur basée sur la compensation d'erreur dans les blocs voisins

5.4.1 La compensation de l'erreur

Comme indiqué dans la section précédente, il est possible de calculer l'erreur Δ' introduite dans un bloc marqué, et estimer l'erreur propagée dans les blocs voisins pour chacun des neuf modes de prédiction possibles (voir chapitre 4, section 4.3).

A l'étape du décodage, la matrice des échantillons d'un bloc A ($\text{Samp}(A)$) est déduite des données résiduelles de A ($\text{Resid}(A)$) et de la prédiction de A ($\text{Pred}(A)$) comme suit:

$$\text{Samp}(A) = \text{Pred}(A) + \text{Resid}(A) \quad (5.9)$$

Lorsque les données secrètes sont insérées dans un bloc et propagées à A , la prédiction est affectée ($\text{NPred}(A)$) et la nouvelle matrice des échantillons de A ($\text{NSamp}(A)$) devient:

$$\begin{aligned}
\text{NSamp}(A) &= \text{NPred}(A) + \text{Resid}(A) \\
\text{NSamp}(A) &= (\text{Pred}(A) + \Delta') + \text{Resid}(A) \\
\text{NSamp}(A) &= \text{NPred}(A) + (\text{Resid}(A) + \Delta') \quad (5.10)
\end{aligned}$$

Il serait possible de préserver les échantillons initiaux de A si les données résiduelles étaient modifiées au niveau du processus du décodage de telle sorte que les nouvelles données résiduelles de A ($\text{NResid}(A)$) deviennent:

$$\text{NResid}(A) = \text{Resid}(A) - \Delta' \quad (5.11)$$

Au niveau du décodeur, les nouveaux échantillons ($\text{N2Samp}(A)$) deviennent:

$$\begin{aligned}
\text{N2Samp}(A) &= \text{NPred}(A) + \text{NResid}(A) \\
\text{N2Samp}(A) &= (\text{Pred}(A) + \Delta') + (\text{Resid}(A) - \Delta') \\
\text{N2Samp}(A) &= \text{Samp}(A) \quad (5.12)
\end{aligned}$$

De ce fait, en soustrayant l'erreur des données résiduelles R (la figure 5.1(a)) au niveau du processus d'encodage, il est possible de réduire la propagation de l'erreur. Toutefois, l'erreur ne peut pas être complètement éliminée car la quantification est une opération irréversible.

5.4.2 Approche proposée : Deuxième schéma

Les données secrètes sont supposées cryptées préalablement pour assurer la sécurité. L'insertion est effectuée dans les blocs de coefficients de luminance DCT 4x4 quantifiés. L'approche proposée peut être décrite en trois parties : Le processus d'insertion, le processus de compensation et le processus de détection.

Dans le processus d'insertion, seulement un coefficient significatif (non nul) est marqué par bloc. Pour appliquer la compensation d'erreur, une matrice initialement nulle M est définie. Son rôle est de sauvegarder les numéros de macroblocs (mb) et les positions des blocs (PositionBlock) qui pourraient être concernés par la compensation d'erreur (blocs voisins tel que représenté à la figure 5.2 (b)). L'algorithme d'insertion est le suivant :

Soit C un coefficient quantifié, $C \neq 0$.

Cas d'insertion du bit 1:

Si $C \bmod 2 = 0$, si $(C > 0)$ $C = C + 1$, sinon $C = C - 1$, $M[\text{mb}][\text{PositionBlock}] = 1$,
Si $C \bmod 2 = 1$, C n'est pas modifié.

Cas d'insertion du bit 0 :

Si $C \bmod 2 = 1$, si $(C > 0)$ $C = C + 1$, sinon $C = C - 1$, $M[\text{mb}][\text{PositionBlock}] = 1$,
Si $C \bmod 2 = 0$, C n'est pas modifié.

Pour le processus de compensation les étapes suivantes sont appliquées:

Etape 1 : L'erreur est calculée pour un bloc marqué et l'erreur propagée ($\Delta'A$) est estimée pour chaque mode de prédiction possible. Par exemple, si le bloc B2 est marqué en ajoutant 1 au

à une amélioration de la qualité de vidéo. Il est clair que la séquence Hall présente la différence de PSNR la plus élevée due aux zones homogènes dominantes. La dégradation peut être vue dans les images de type P, pour $N > 5$. La figure 5.7 présente une image de type P (numéro 33) marquée et non marquée de la séquence Hall.

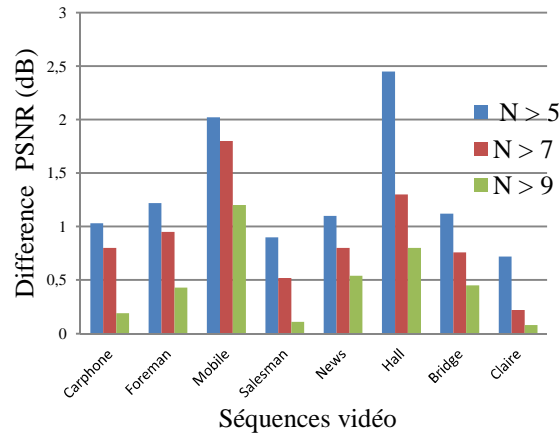


Figure 5.6 : Les différences de PSNR calculées pour 8 séquences pour $N > 5$, $N > 7$ et $N > 9$.

Les résultats montrent que le débit est maintenu pour toutes les séquences et pour toutes les valeurs de test de N comme représenté à la figure 5.8.

La figure 5.9 montre que la capacité d'insertion réduit avec l'augmentation de N . La séquence Mobile présente une capacité d'insertion élevée liée à la texture détaillée de la vidéo.

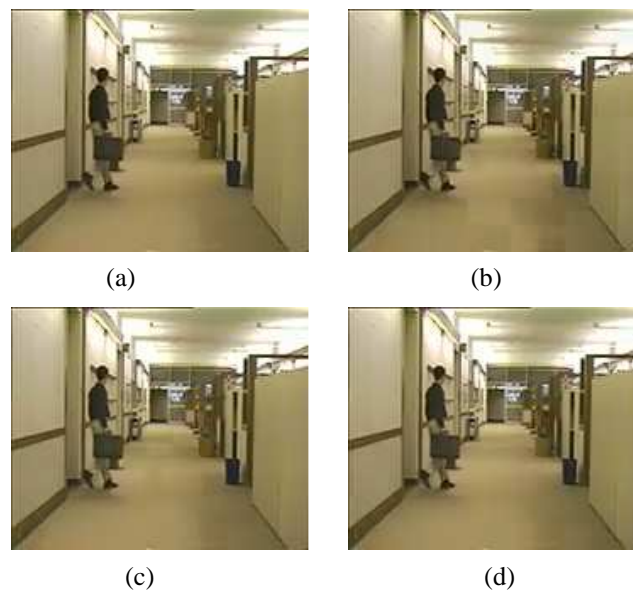


Figure 5.7 : La séquence Hall, image P n° 33 (a) image non marquée, (b) image marquée ($N > 5$), (c) image marquée ($N > 7$), (d) image marquée ($N > 9$).

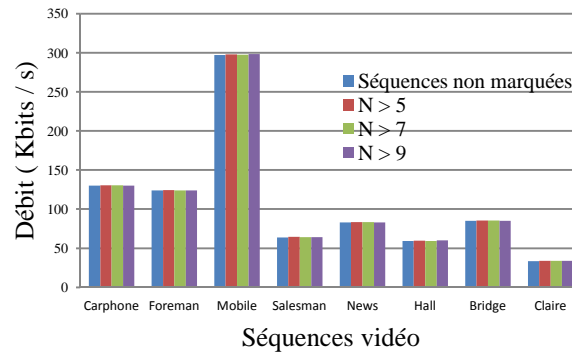


Figure 5.8 : Le débit de 8 séquences marquées ou non pour $N > 5$, $N > 7$ et $N > 9$.

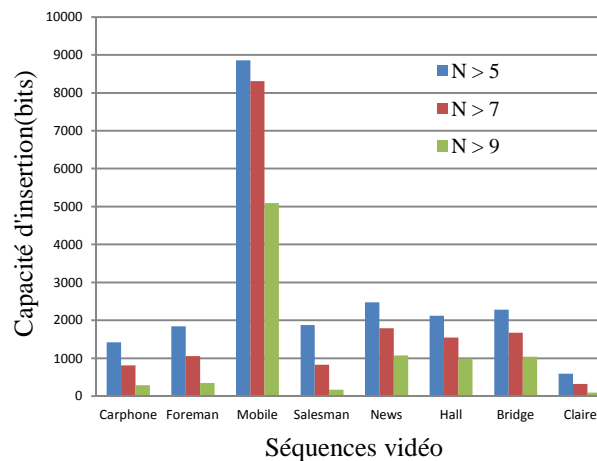


Figure 5.9 : Le débit de 8 séquences marquées ou non pour $N > 5$, $N > 7$ et $N > 9$.

5.5.2 Deuxième schéma

Dans le but de choisir des coefficients qui contiendraient les bits secrets dans un bloc DCT 4x4, des tests ont été effectués pour différentes positions de coefficients non nuls (pour chaque test la position est la même dans chaque bloc). Les résultats ont montré qu'il était possible d'obtenir une amélioration importante de la qualité de l'image si les blocs voisins à un bloc 4x4 marqué n'étaient pas eux même marqués. Dans le tableau 5.4 sont enregistrés les résultats de l'augmentation du débit et de la capacité d'insertion pour le marquage d'un coefficient DC dans chaque bloc 8x8 (4 coefficients dans le macrobloc). La qualité visuelle est mesurée par la différence de PSNR entre une séquence marquée et non marquée. Nous présentons à la figure 5.10 l'apport de la compensation de l'erreur sur la qualité.

L'approche proposée offre de bons résultats dans l'amélioration de la qualité de la vidéo tel que représenté sur la figure 5.10. Un exemple est donné dans la figure 5.11 pour la première image de la séquence Mother (la correction de la dégradation est visible particulièrement sur la joue de la mère). Le débit est maintenu et même réduit tel que pour le cas de la séquence Mother qui présente aussi la capacité d'insertion la plus élevée. En effet, nos expérimentations montrent que

la texture (riche en détails ou homogène) n'a pas une grande influence sur la capacité d'insertion car l'insertion est effectuée dans le coefficient DC.

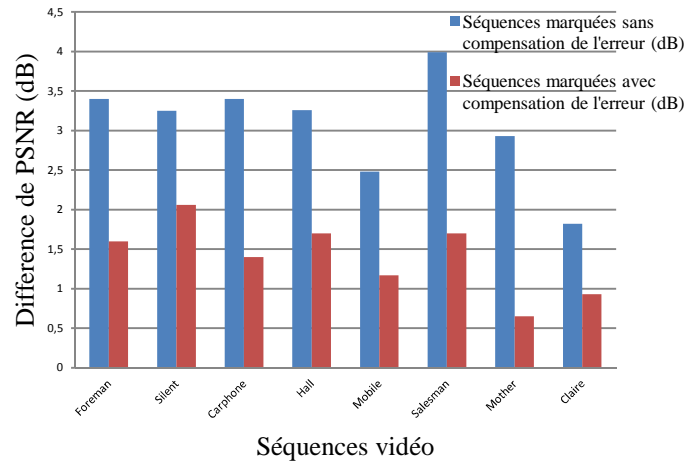


Figure 5.10 : Différences de PSNR mesurées pour huit séquences.

Tableau 5.4. Augmentation du débit et de la capacité l'insertion pour huit séquences vidéo.

Séquences vidéo	Augmentation de débit (%)	Capacité d'insertion (bits)
Foreman	0.06	1815
Silent	0.15	2315
Carphone	0.41	1663
Hall	0,32	1676
Mobile	-0.03	2852
Salesman	0.79	2234
Mother	-0.12	2913
Claire	0.26	782



(a)



(b)



(c)

Figure 5.11 : Séquence Mother, (a) Image non marquée, (b) Image marquée sans correction d'erreur, (c) Image marquée avec correction d'erreur.

5.5.3 Comparaison avec d'autres méthodes

Dans le tableau 5.5, une comparaison est présentée entre les approches proposées et les résultats présentés dans [9] pour quatre séquences vidéo. L'implémentation est faite dans les mêmes conditions.

Dans le premier schéma, les résultats sont donnés pour N égal à 7, quant au deuxième schéma, les résultats sont donnés pour une insertion de deux bits par macrobloc seulement.

Nous remarquons, pour les deux schémas que les approches proposées présentent de meilleurs résultats en termes de qualité de la vidéo, capacités d'insertion et particulièrement pour la préservation du débit, ce qui est principalement dû au fait que les coefficients nuls ne sont pas utilisés pour l'insertion des données secrètes. A l'opposé, l'algorithme proposé dans [9] peut utiliser les coefficients nuls, sinon la capacité d'insertion serait très faible, car il ne sera pas évident de rencontrer fréquemment les conditions d'insertion comme mentionné dans la section 5.3.1.

Tableau 5.5. Comparaison entre nos approches et celle de Ma.

Séquences		Algorithme de Ma	Approche proposée (schema 1)	Approche proposée (scheme2)
Carphone	Capacité moy. dans I (bits)	806	816	844
	Difference PSNR (dB)	0.93	0.8	0.76
	Augmentation du débit (%)	3.24	0.35	0.35
Foreman	Capacité moy. dans I (bits)	737	1059	904
	Difference PSNR (dB)	0.91	0.95	0.75
	Augmentation du débit (%)	3.16	-0.24	0.28
Mobile	Capacité moy. dans I (bits)	910	8311	1490
	Difference PSNR (dB)	0.72	1.8	0.92
	Augmentation du débit (%)	0.88	0.16	-0.04
Salesman	Capacité moy. dans I (bits)	936	833	1155
	Difference PSNR (dB)	1.78	0.52	0.73
	Augmentation du débit (%)	7.17	0.2	0.28

Le tableau 5.6 présente une comparaison entre les résultats présentés dans [72] et nos approches implémentées dans les mêmes conditions. Pour le premier schéma, la capacité d'insertion est liée à la disponibilité des paires de coefficients appariés, par contre, il est possible d'utiliser plus de coefficients par macrobloc dans le second schéma.

Nous remarquons que le second schéma donne de meilleurs résultats que le premier en termes de capacité d'insertion. Cependant, en comparant à [72], les deux approches proposées semblent s'adapter plus aux séquences riches en détails (telle la séquence Mobile) pour lesquels le compromis entre les critères de tatouage est mieux atteint. En effet, pour cette séquence, la capacité d'insertion peut être augmentée tout en maintenant relativement la qualité vidéo. Le débit binaire est plus élevé mais reste inférieur à 1%. De plus, les méthodes proposées présentent l'avantage d'avoir des marques «à extraire» et elles peuvent être implémentées au cours du processus d'encodage pour la diffusion en temps réel.

Tableau 5.6. Comparaison entre nos approches et celle de Huo.

Sequences		Algorithme de Huo	Approche proposée (schema 1)	Approche proposée (schema 2)
		à détecter	à extraire	
Mobile	Capacité moy. dans I (bits)	672	700	870
	différence PSNR (dB)	38.61	37.65	37.25
	Augmentation du débit (%)	0.40	0.80	0.84
Foreman	Capacité moy. dans I (bits)	428	218	400
	différence PSNR (dB)	40.58	42.43	36.83
	Augmentation du débit (%)	1.28	1.26	1.29

5.6 Conclusion

Dans ce chapitre, nous avons proposé de nouvelles approches ayant pour objectif de limiter la propagation de l'erreur introduite lors du marquage du codec H.264/AVC.

Ces méthodes sont basées sur des mesures préalables du signal d'erreur introduit dans un bloc (lors du marquage d'une paire de coefficients) ou bien dans les blocs voisins (lors du marquage d'un seul coefficient).

Les deux schémas sont appliqués durant le processus d'encodage, ils sont convenables pour la diffusion de vidéo en temps réel, et ils présentent une capacité d'insertion relativement élevée pour une baisse négligeable du PSNR, et un débit préservé, particulièrement pour les séquences hautement texturées.

Comme perspective pour le premier schéma, il serait intéressant d'investiguer la possibilité de collecter des informations durant le processus d'encodage sur les blocs qui n'ont pas été codés, ce qui pourrait aider dans le choix des paires de coefficients à marquer.

Le second schéma conduit également à des perspectives intéressantes liées à la correction de la propagation de la distorsion lorsque les données secrètes sont insérées dans les coefficients DC, ce qui pourrait améliorer la robustesse des méthodes de tatouage basées sur l'insertion de données dans les basses fréquences.

Conclusions et perspectives

L'objectif de notre travail était d'étudier le codec vidéo H.264/AVC et d'explorer de nouvelles possibilités d'insérer des données de manière secrète pouvant satisfaire certaines applications telles que : le contrôle d'authentification, la communication secrète, l'archivage de dossier médical, etc. La difficulté résidait essentiellement dans la nécessité d'améliorer le compromis entre les différentes caractéristiques du tatouage, telles que la capacité d'insertion, le débit, la qualité de la vidéo, etc.

Dans un premier temps, nous nous sommes intéressées aux méthodes de tatouage réversible qui permettaient de revenir au document initial. Au niveau du codec, le module de transformation et de quantification qui effectue le calcul des coefficients quantifiés est utilisé pour insérer des bits secrets dans les bits les moins significatifs (LSBs) des coefficients.

Les deux schémas proposés sont l'adaptation des méthodes proposées initialement pour des images. La première méthode est basée sur la technique du décalage de l'histogramme appliquée aux images de types I et P. Les résultats ont montré qu'il était possible d'appliquer la méthode du décalage de l'histogramme au codec H.264/AVC, et selon les besoins d'une application donnée en termes de capacité d'insertion, de débit et de qualité visuelle, il était possible de choisir le coefficient adéquat pour l'insertion, ainsi que le type d'image. Quant à la deuxième méthode, elle est basée sur l'exploitation des ensembles de coefficients nuls qui faciliterait la reconnaissance du coefficient qui a été modifié et la récupération de la vidéo initiale. Les règles de mise en correspondance qui ont été utilisées ont permis d'améliorer la capacité d'insertion et la qualité visuelle.

Nous nous sommes intéressées ensuite aux méthodes basées sur les modes de prédiction. En se basant sur la direction des modes de prédiction, nous avons proposé un schéma qui ne nécessite pas des connaissances préalables des différents éléments du codage, ce qui faciliterait son application dans le cas des diffusions en temps réel.

Nous avons présenté en dernier lieu deux méthodes pour la réduction de la propagation d'erreur causées par l'insertion des bits secrets dans les coefficients quantifiés. Les résultats ont montré qu'en effectuant une compensation de l'erreur dans le bloc marqué ou dans les blocs voisins, il était possible d'apporter une amélioration considérable à la qualité de la vidéo tout en maintenant la capacité de l'insertion et le débit de la vidéo.

Enfin, nous constatons que malgré la complexité du codec H.264/AVC, il a été possible d'exploiter ses différents modules pour insérer des données supplémentaires, et diverses perspectives de ce travail peuvent être envisagées, parmi lesquelles :

- L'amélioration des caractéristiques des méthodes présentées, par rapport à des applications précises.
- L'exploitation du module du codage entropique, du flux vidéo, et différentes configurations des profils pour l'insertion des données.
- L'introduction des méthodes réversibles au niveau des modes intra pour tirer profit du bon compromis qui existe entre les critères : capacité d'insertion, qualité et débit de la vidéo.
- L'évaluation et l'amélioration des méthodes proposées en termes de robustesse étant donné l'indispensabilité de ce critère pour certaines applications telle que la protection des droits d'auteurs.

Bibliographie

- [1] S. Bouchama, H. Aliane, H. Boudier, Les moyens techniques de protection des droits d'auteurs : Apports du tatouage, Séminaire national sur Le cadre juridique des TIC en Algérie : Entre opportunités et contraintes ? Alger (Algérie), Mai 2012.
- [2] T. Wiegand, G.J. Sullivan, G. Bjontegaard, and A. Luthra, Overview of the H.264/AVC video coding standard, *IEEE Trans. Circuits Syst. Video Technol.*, vol.13, n°7, July 2003.
- [3] I.E.G. Richardson, H264 and MPEG4 video compression: Video coding for next-generation multimedia, The Robert Gordon University, Aberdeen, UK, Wiley, 2003.
- [4] M. Noorkami and R.M. Mersereau, Compressed domain video watermarking for H264, in *Proc. of IEEE International Conference of Image Processing*, Genoa (Italy), 2005.
- [5] A. Golikeri, P.Nasiopoulos, and Z. J.Wang, An improved scalar quantization-based digital video watermarking scheme for H.264/AVG. *IEEE International Symposium on Circuits and Systems*, 2006.
- [6] G. Yang, J. Li, Y. He, and Z. Kang, An information hiding algorithm based on intra-prediction modes and matrix coding for H.264/AVC video stream, *Int. J. Electron. Commun*, vol.64, pp.331-337, 2010.
- [7] Y. Hu, C. Zhang, and Y. Su, Information hiding based on intra prediction modes for H.264/AVC, *IEEE International Conference on Multimedia and Expo (ICME)*, pp.1231-1234, 2007.
- [8] D. Xu, R. Wang, and J. Wang, Prediction mode modulated data-hiding algorithm for H.264/AVC, *J Real-Time Image Proc.* vol.5, 2010.
- [9] X. Ma, Z. Li, H. Tu, and B. Zhang, A data hiding algorithm for H.264/AVC video streams without intra-frame distortion drift, in *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, n°10, 2010.
- [10] M. Ghanbari, *Standard codecs: Image compression to advanced video coding*, IEE Press, June 2003.
- [11] C. Alexandre and H.V. Thien, Algorithme rapide d'estimation de mouvement, *Traitement du Signal*, vol.13, n° 4, 1996.
- [12] J-N. Gouyet and F. Mahieu, MPEG-4 : Advanced video coding: Systèmes et applications, *Techniques de l'ingénieur*, vol.TEB2, n° TE5367, 2007.
- [13] JVT-ISO/IEC 14496- 10 AVC- ITU-T Recommendation H264, Advanced video coding for generic audio visual services, Draft ITU-T Recommendation and Final draft international standard, JVT- G050r1, 2005.
- [14] D. Marpe, H. Schwarz, and T. Wiegand, Context-based adaptive binary arithmetic coding in the H.264/AVC video compression standard, *IEEE Trans. Circuits Syst. Video Technol.*, vol.13, n°7, pp.620-636, 2003.
- [15] G.J. Sullivan, P. Topiwala, and A. Luthra, The H.264/AVC Advanced video coding standard: Overview and introduction to the fidelity range extensions, in *Proc. of the SPIE Conference on applications of digital image processing 7 XXVII (Special Session on Advances in the Emerging H.264/AVC Video Coding Standard)*, vol.5558, n°1, pp.454-474, 2004.
- [16] L. Cieplinski, Scalable video coding for flexible multimedia services, *IEEE Trans. Circuits Syst. Video Technol.*, vol.18, n°7, 2006.
- [17] G. Doerr, and J.L. Dugelay, A guide tour of video watermarking, *Signal process-image*, vol.18, pp.263-282, April 2003.

- [18] K. Tanaka, Y. Nakamura, and K. Matsui, Embedding secret information into a dithered multi-level image, in Proc. of IEEE Military Communications Conference, vol. 1, Monterey CA (USA), pp. 216–220, October 1990.
- [19] A.Z. Tirkel, G.A. Rankin, R.M.V. Schyndel, W.J. Ho, N. Mee, and C.F. Osborne, Electronic watermark, Digital Image Computing Technology and Applications (DICTA'93). Sidney, pp. 666–673, 1993.
- [20] W. Bender, D. Gruhl, and N. Morimoto, Techniques for data hiding, in Proc. of the SPIE - The International Society for Optical Engineering, vol.2420, San Jose CA (USA), pp.164–173, February 1995.
- [21] A.Z. Tirkel, R.G.V. Schyndel, and C.F. Osborne, A two dimensional digital watermark, in Digital Image Computing, Techniques and Applications (DICTA'95), Brisbane Australia, pp.378–383, December 1995.
- [22] I.J. Cox, J. Kilian, and T. Shamoan, Secure spread spectrum communication for multimedia. Technical report, NEC Research Institute, Princeton NJ (USA), 1995.
- [23] M. Kutter, F. Jordan, and F. Bossen, Digital signature of color images using amplitude modulation, in Proc. of the SPIE- International Society for Optical Engineering, vol. 3022, pp.518–526, USA, February 1997.
- [24] B. Chen, and G.W. Wornell, Quantization index modulation: A class of provably good methods for digital watermarking and information embedding, IEEE Trans. Inform. Theory, vol. 47, pp. 1423–1443, May 2001.
- [25] I.J. Cox, J. Kilian, F.T. Leighton, and T. Shamoan, Secure spread spectrum watermarking for multimedia, IEEE Trans. Image Processing, vol. 6, pp. 1673–1687, December 1997.
- [26] H. Park, S.H. Lee, and Y.S. Moon, Adaptive video watermarking utilizing video characteristics in 3D-DCT domain, in Proc. of the 5th International Workshop on Digital Watermarking (IWDW), Korea, pp.397-406, 2006.
- [27] G. Doërr and J.L. Dugelay, Collusion issue in video watermarking, in Proc. of the SPIE Conference on Security, Steganography and Watermarking of Multimedia Contents VII, vol.5681, pp. 685-696, 2005.
- [28] J. Dittmann, M. Stabenau, and R. Steinmetz, Robust MPEG video watermarking technologies, ACM Multimedia 98, ACM Press, New York, pp. 71-80, 1998.
- [29] M. Kutter, F. Jordan, and T.Ebrahimi, Proposal of watermarking technique for hiding/retrieving data in compressed and decompressed video, Technical report M2281, ISO/IEC document, JTC1/SC29/WG11, 1997.
- [30] K.W. Kang, K.S. Moon, G.S. Jung, and J. N. Kim, An efficient video watermarking scheme using adaptive threshold and minimum modification on motion vectors, in Proc. of International Conference on Image Analysis and Recognition (ICIAR), pp.294-301, 2005.
- [31] T.Y. Kuo, Y.C. Lo, and C.I. Lin, Fragile video watermarking technique by motion field embedding with rate-distortion minimization, in Proc. of IEEE International Conference on Intelligent Information Hiding and Multimedia Signal Processing, pp.853-856, 2008.
- [32] T.Y. Chen, T.H. Chen, Y.T. Lin, Y.C. Chang, and D.J.Wang, H.264 video authentication based on semi-fragile watermarking, Int. J. of Innov. Comput. I., vol.6, n° 3(B), pp.1411-1420, 2010.
- [33] M.A. Ali, E.A. Edirisinghe, A semi-fragile watermarking technique for H.264/AVC using CAVLC. Int. J of Signal and Image Proc., vol.1, pp.151-159, 2010.
- [34] S. Bouchama, L. Hamami, H. Aliane, Watermarking techniques applied to H264/AVC video standard, in proc. of International conference on information science and applications (ICISA 2010), Seoul (Korea), April 2010.

- [35] W.M. Chen, C.J. Lai, and C.C.Chang, H.264 video watermarking with secret image sharing, IEEE International Symposium on Broadband Multimedia Systems and Broadcasting, Bilbao (Spain), May 2009.
- [36] K. Ait Saadi, A. Bouridane, and H. Meraoubi, Secure and robust copyright for H.264/AVC based on selected blocks DCT, in Proc. of International Conference on Signal Processing and Multimedia Applications, pp.351-355, 2008.
- [37] Y.H. Seo, H.J. Choi, C.Y. Lee, and D.W. Kim, Low-complexity watermarking based on entropy coding in H.264/AVC, IEICE Trans. on Fund. of Electron Comm. Comput. Sci., pp. 2130-2137, 2008.
- [38] S.K. Kapotas and A.N. Skodras, Real time data hiding by exploiting the IPCM macroblocks in H.264/AVC streams, J. Real-Time Image Proc., vol.4, pp.33-41, 2009.
- [39] D. Pröfrock, H. Richter, M. Schlauweg, and E. Müller, H.264/AVC video authentication using skipped macroblocks for an erasable watermark, in Proc. of the IEEE Conference on Visual communication and Image Processing (VCIP), Beijing (China), 2005.
- [40] R. Iqbal, S. Shirmohammadi, and J. Zhao, Hard authentication of H.264 video applying MPEG-21 generic bitstream syntax description (GBSD), in Proc. of IEEE International Conference on Multimedia & Expo, Beijing (China), 2007.
- [41] S. Bouchama, L. Hamami, T.Qadri and M. Ghanbari. Watermarking of H.264 coded video based on the shifted-histogram technique, International journal of latest trends in computing, vol.1, n° 2, pp.109-115, 2010.
- [42] M.A. Ali, E.A. Ediriginshe, High-Capacity Reversible Multi-Layer Watermarking for H.264/AVC Video, 7th International Conference on Computing and Convergence Technology (ICCCT), pp. 631-636, Seoul (Korea), 2012.
- [43] J. Tian, Reversible data embedding using a difference expansion, IEEE Trans. Circuits Syst. Video Technol., vol. 13, n° 8, 2003.
- [44] S .Bouchama, L. Hamami, H. Aliane, Watermarking of compressed video based on DCT coefficients and watermark preprocessing, International Joint Conference on Computer, Imaging and Computer Graphics Theory and Applications, Vilamoura (Portugal), 2011.
- [45] A. Punchihewa, D.G. Bailey, R.M. Hodgson, A survey of coded image and video quality assessment, The 28th International Conference on Image and Vision Computing, New Zealand, November 2003.
- [46] M. Yuen and H.R. Wu, A survey of hybrid MC/DPCM/DCT video coding distortions, Signal processing (EURASIP), vol.70, pp.247-278, 1998
- [47] BT.500-11 (UIT), Methodology for the subjective assesement of the quality of television pictures, ITU recommendations, 2002.
- [48] M.J. Nadenau, S.Winkler, D. Alleysson, and M.Kunt, Human vision models for perceptually optimized image processing – A review, IEEE Trans. Image processing, 2003.
- [49] S. Winkler, Vision models ans quality metrics for image processing applications, Thèse de doctorat, Ecole Polytechnique Fédérale de Lausanne, 2000.
- [50] H. H.Y. Tong and A.N. Venetsanopoulos, A perceptual model for JPEG application on block classification, texture masking, and luminance masking, in Proc. of International Conference on Image Processing (ICIP), pp. 428-432, 1998.
- [51] K.T. Tan, and M. Ghanbari, Blockiness detection for MPEG2-coded video, IEEE Signal Processing Lett., vol.7, n° 8, pp. 213-215, 2000.

- [52] I.P. Gunawan, and M. Ghanbari, Reduced-reference video quality assessment using discriminative local harmonic strength with motion, *IEEE Trans. Circuits Syst. Video Technol.*, vol.18, n°.1, pp. 71-73, 2008.
- [53] M.G. Choi, J.H. Jung, and J.W. Jeon, No-reference image quality assessment using blur and noise, *Int. J. Electr. Electron. Eng.*, vol.3, n°.3, pp.18-22, 2009.
- [54] H. Liu, N. Klomp, and I. Heynderickx, A no-reference metric for perceived ringing artifacts in images, *IEEE Trans. Circuits Syst. Video Technol.* vol.20, n°.4, April 2010
- [55] O.M. Al-Qershi, ROI-based tamper detection and recovery for medical images using reversible watermarking technique, in *Proc. of IEEE International Conference on Information Theory and Information Security (ICITIS)*, pp.151-155, Beijing (China), 2010.
- [56] A.S. Brar and M. Kaur, A survey of reversible watermarking techniques for data hiding with ROI-tamper detection in medical images, in *Proc. of Communications in Computer and Information Science*, vol.296, pp.516-522, 2013.
- [57] S. Bouchama, L. Hamami, H. Aliane, Reversible data hiding scheme for the H.264/AVC codec, in *proc. of the 4th International conference on information science and technology (ICISA2013)*, Pattaya (Thailand), June 2013.
- [50] M. Chaumont and W. Puech, A high capacity reversible watermarking scheme, in *Proc. of the SPIE Conference on Visual Communications and Image Processing*, vol.7257, 2009.
- [59] R. Caldelli, F. Filippini, and R. Becarelli, Reversible watermarking techniques: An overview and a classification, *EURASIP J. Inform. Secur.*, 2010.
- [60] J-B. Feng, I-C. Lin, C-S. Tsai, and Y-P. Chu, Reversible watermarking: current status and key issues, *Int. J. Network Security*, vol.2, no.3, pp.161–171, May 2006.
- [61] Z. Ni, Y.Q. Shi, N. Ansari, and W. Su, Reversible data hiding, *IEEE Trans. Circuits Syst. Video Technol.*, vol.16, n°.3, pp.354-362, 2006.
- [62] D. Coltuc and J. M. Chassery, Very fast watermarking by reversible contrast mapping, *IEEE Signal Processing Lett.*, vol. 14, n°. 4, pp. 255–258, April 2007.
- [63] R.A. Farrugia, Reversible visible watermarking for H.264/AVC encoded video, in *Proc. IEEE of International Conference on Computer as a Tool (EUROCON)*, Lisbon (Portugal), April 2011.
- [64] C-C. Chang, C-C. Lin, C-S. Tseng, and W-L. Tai, Reversible hiding in DCT-based compressed images, *Int. J. Info. Sci.*, vol.177, n°.13, pp. 2768–2786, July 2007.
- [65] S. Gujjunoori and B.B. Amberker, A DCT Based Reversible Data Embedding Scheme for MPEG-4 Video Using HVS Characteristics, in *Proc. of the 8th Indian Conference on Computer Vision, Graphics and Image processing (ICVGIP)*, 2012.
- [66] M. Fallahpour, D. Megias and M. Ghanbari, High capacity, reversible data hiding in medical images, *IET Image Proc.*, vol.5, n°.2, pp.190-197, 2011.
- [67] S. Bouchama, L. Hamami, H. Aliane, H.264/AVC data hiding based on intra prediction modes for real time applications, in *Proc. of International conference on signal processing and imaging engineering (ICSPIE'12)*, San Francisco (USA), October 2012.
- [68] P. Wang, Z. Zheng, and L. Li, A video watermarking scheme based on motion vectors and mode selection, in *Proc. of IEEE International Conference on Computer Science and Software Engineering*, pp.233-237, 2008.
- [69] S. Bouchama, L. Hamami, H. Aliane, A comparative study between compressed video watermarking methods based on DCT coefficients and intra prediction, *Séminaire Inter-laboratoire des doctorants en informatique*, Constantine (Algérie), Avril 2011.

- [70] S. Bouchama, L. Hamami, H. Aliane, H.264/AVC data hiding algorithm with a limitation of the drift distortion, in proc. of the 11th International conference on signal processing, Beijing (China), October 2012.
- [71] X. Gong and H-M. Lu, Towards fast and robust watermarking scheme for H.264 video, in Proc. 10th IEEE International Symposium on Multimedia (ISM), pp.649-653, 2008.
- [72] L. Zhang, Y. Zhu, and L-M. Po, A novel watermarking scheme with compensation in bit-stream domain for H.264/AVC, in Proc. IEEE International Conference on Acoustic Speech and Signal Processing (ICASSP), pp.1758–1761, Dallas TX (USA), 2010.
- [73] W. Huo, Y. Zhu, and H. Chen, A Controllable Error-Drift Elimination Scheme for Watermarking Algorithm in H.264/AVC Stream, IEEE Signal Processing. Lett., vol. 18, n° 9, pp.535-538, 2011.
- [74] S. Bouchama, L. Hamami, H. Aliane, Error drift compensation for data hiding of the H.264/AVC, Journal of control engineering and applied informatics, vol.15, n°3 pp. 70-77, 2013.

Annexe 1: Contributions de l'auteur

• Journaux Internationaux

1. S. Bouchama, L. Hamami, H. Aliane, Error drift compensation for data hiding of the H.264/AVC, Journal of control engineering and applied informatics, vol.15, n°3 pp. 70-77, 2013.
2. S. Bouchama, L. Hamami, T. Qadri and M. Ghanbari, Watermarking of H.264 coded video based on the shifted-histogram technique, International journal of latest trends in computing, vol.1, n° 2, pp.109-115, 2010.

• Conférences Internationales

3. S. Bouchama, L. Hamami, H. Aliane, Reversible data hiding scheme for the H.264/AVC codec, in proc. of the 4th International conference on information science and technology (ICISA2013), Pattaya (Thailand), June 2013.
4. S. Bouchama, L. Hamami, H. Aliane, H.264/AVC data hiding algorithm with a limitation of the drift distortion, in proc. of the 11th International conference on signal processing, Beijing (China), October 2012.
5. S. Bouchama, L. Hamami, H. Aliane, H.264/AVC data hiding based on intra prediction modes for real time applications, in Proc. of International conference on signal processing and imaging engineering (ICSPIE'12), San Francisco (USA), October 2012.
6. S. Bouchama, L. Hamami, H. Aliane, Watermarking of compressed video based on DCT coefficient and watermark preprocessing, International Joint Conference on Computer, Imaging and Computer Graphics Theory and Applications, Vilamoura (Portugal), 2011.
7. S. Bouchama, L. Hamami, H. Aliane, Watermarking techniques applied to H264/AVC video standard, in proc. of International conference on information science and applications (ICISA 2010), Seoul (Korea), April 2010.
8. S. Bouchama , L. Hamami, Les techniques de tatouage appliquées à l'imagerie médicale, Les journées d'études Algéro-françaises sur l'imagerie médicale (Jetim08), Tipaza (Algérie), 2008.
9. S. Bouchama , L. Hamami, Le tatouage pour l'intégrité et la confidentialité des données médicales, 2nd International conference on electrical and electronics engineering, Laghouat (Algérie), 2008.

• Conférences Nationales

10. S. Bouchama, H. Aliane, H. Boudier, Les moyens techniques de protection des droits d'auteurs : Apports du tatouage, Séminaire national sur «Le cadre juridique des TIC en Algérie : Entre opportunités et contraintes ? Alger (Algérie), Mai 2012.
11. S. Bouchama, L. Hamami, H. Aliane, A comparative study between compressed video watermarking methods based on DCT coefficients and intra prediction, Séminaire Inter-laboratoire des doctorants en informatique, Constantine (Algérie), Avril 2011.

Annexe 2: Glossaire français-anglais

A

Appariement de bloc (méthode par)	block matching
--------------------------------------	----------------

B

Base (profil de)	Baseline profile
Brouillage	Scrambling

C

Capacité d'insertion	Embedding capacity
----------------------	--------------------

D

Débit	bitrate
-------	---------

E

Empreinte	Fingerprint
Etendu (profil)	Extended profile
Evolutif (profil)	Scalable profile

F

Flux (vidéo)	Stream
--------------	--------

G

Gigue (attaque de)	Jitter attack
---------------------	---------------

I

Image	Image, frame.
Inversibilité (attaque d')	Deadlock attack

M

Macrobloc	Macroblock
Modélisation de contexte	Context modeling

P

Principal (profil)	Main profile
--------------------	--------------

R

Résidus (données résiduelles)	Residual data
-------------------------------	---------------

T

Table de correspondance	Look-up table
Tranche	Slice

V

Séquence vidéo	Video sequence
Vecteur de mouvement	Motion vector
