



Année :
N° d'ordre :
N° de série :

THÈSE

EN VUE DE L'OBTENTION DU DIPLOME DE
DOCTORAT 3^{ème} CYCLE
DE L'UNIVERSITE DES FRERES MENTOURI, CONSTANTINE

Filière : ELECTRONIQUE
Option : Automatique et Traitement du Signal

Présentée par
BOUDECHICHE Djamel Eddine

Thème

**Codage vidéo distribué pour les images médicales
application aux capsules endoscopiques**

Soutenue le :26/02/2018

Devant le Jury composé de :

Nom et Prénom	Grade		
Mr BENNIA Abdelhak	Prof	Univ. de Constantine 1	Président
Mr KHAMADJA Mohammed	Prof	Univ. de Oum El Bouaghi	Rapporteur
Mme HACHOUF Fella	Prof	Univ. de Constantine 1	Examineur
Mr BAATOUCHE M.Chawki.	Prof	Univ. de Constantine 2	Examineur
Mr FERDI Youcef	Prof	Ecole Biotechnologie	Examineur
Mr BENIERBAH Said	MCA	Univ. de de Constantine 1	Invité

Remerciements

Mon travail est réalisé en vue d'obtention du diplôme du Doctorat en électronique. Au terme de ce travail nos vifs remerciements vont à:

Mon promoteur **Pr. Khamadja Mohammed** pour avoir proposé et en charge ce projet, il a mit un soin particulier à lire, à discuter et à corriger mon mémoire. Sans oublier ses précieux conseils, sa patience et sa simplicité pour tous les efforts qu'il a fournis, pour la confiance qu'il m'a accordé. Je lui exprime ma profonde gratitude.

Dr. Benierbah Said pour son aide sur le plan technique et les grandes qualités humaines qui m'ont permis de mener à bout ce travail. Son soutien s'est avéré déterminant pour mener ce travail à terme.

Je remercie vivement Prof. **Abdelhak BENNIA** pour avoir accepté de présider ce jury et juger ce travail.

Je tiens à remercier les membres de jury, les professeurs **Fella HACHOUF**, **M.Chawki BAATOUCHE** et **Youcef FERDI**, pour l'attention qu'ils ont portée à ce travail. Je leur suis très reconnaissant d'avoir accepté de juger mon travail en me faisant l'honneur d'être mes examinateurs de cette thèse.

Enfin, j'espère que mon travail témoigne de ma bonne volonté, mon assiduité et mon sérieux.

DEDICACE

Je dédie ce travail à :

Mon cher pays avec la ferme volonté de servir aussi pleinement que possible.

Mon père à qui je dois beaucoup et qu'est pour moi la source d'espoir, du savoir, du courage et de la patience.

Ma très chère maman qui a été toujours pour moi une source de tendresse et de courage.

Mon frère BILLEL, et mes sœurs RIMA et SYRINIA.

Mes ancles et tantes, mes cousins et cousines.

Tous mes amis

Et à toute la promotion d'électronique.

Résumé

Le codage vidéo distribué (DVC) est un nouveau paradigme de la compression qui permet de réaliser un codage moins complexe, robuste et avec moins de latence pour certaines applications. Le codage vidéo traditionnel contient un codeur de 5 à 10 fois plus complexe que le décodeur et nécessite plus de ressources énergétiques et matérielles pour le codage. Cependant, le codage distribué est un candidat pour les applications sans fil qui contiennent des sources énergétiques limitées. La capsule endoscopique sans fil est un outil médical présenté en 2000 et utilisé par les médecins dans le diagnostic des maladies gastro-intestinal. Cet outil ne peut visualiser tout le système gastro-intestinal, car le temps de vie de batterie est limité. L'une des solutions proposées est d'utiliser un système de compression simple et efficace pour réduire le nombre de bits transmis par l'antenne, qui consomme la majorité de l'énergie de la batterie. Dans cette thèse, nous proposons de nouveaux systèmes de codage vidéo distribué simple, efficace et applicable pour les images de la capsule endoscopique. Il s'agit des systèmes de DVC simple qui utilise un seul type de trame contrairement aux systèmes traditionnels de DVC qui utilisent deux types de trames. Les deux systèmes proposés de DVC utilisent la base de données pour générer l'information adjacente (SI) qu'elle utilise pour décoder la trame source. La sélection dans la base de données est effectuée en utilisant les coefficients DC dans la 1^{re} contribution et l'image sous-échantillonnée dans la seconde. Pour la troisième contribution, le codec utilise la quantification vectorielle pour générer l'information adjacente (SI). Plusieurs algorithmes sont implémentés dans les systèmes proposés pour améliorer l'efficacité de la compression. Les résultats obtenus sont significatifs en comparaison avec le codec de la Motion JPEG déjà utilisée dans la capsule endoscopique.

Mots clés : le codage vidéo distribué, capsule endoscopique, la base des données, la quantification vectorielle, l'information adjacente.

Abstract

Distributed video coding (DVC) is a new compression paradigm that enables less complex, robust and less latency encoding for some applications. Traditional video coding contains an encoder 5 to 10 times more complex than the decoder and requires more energy and hardware resources for coding. However, distributed coding is a candidate for wireless applications that contain limited energy sources. The wireless capsule endoscopy is a medical tool introduced in 2000 and used by doctors in the diagnosis of gastrointestinal diseases. This tool cannot visualize the completely gastrointestinal system because the battery life is limited. One of the proposed solutions is to use a simple and efficient compression system to reduce the number of bits transmitted by the antenna that consume the most energy from the battery. In this thesis, we propose new distributed video coding systems simple, efficient and applicable for the images of the capsule endoscopy. The proposed DVC systems are a low complexity because uses a single type of frame compared to traditional DVC systems that use two types of frames. Both proposed DVC systems use the database to generate the side information (SI) that used to decode the source frame. The selection made in the database using the coefficients DC in the first contribution and the subsample image in the second. For the third contribution, the codec uses the vector quantization to generate the side information (SI). Several algorithms are implemented in the proposed systems to improve compression efficiency. The results obtained are significant and prove the ideas proposed by comparing them with the Motion JPEG codec already used in the endoscopic capsule.

Keywords : Distributed video coding, capsule endoscopy, databases, vector quantization, side information.

الملخص

ان ترميز الفيديو الموزع (DVC) هو نموذج ضغط جديد يتيح ترميز أقل تعقيدا و قوة و أقل كمون لبعض التطبيقات و يحتوي الترميز الفيديو التقليدي على مشفر أكثر تعقيدا من 5 الى 10 مرات من جهاز فك الترميز ويتطلب المزيد من موارد الطاقة و الأجهزة اللازمة للترميز. غير أن الترميز الموزع مرشح قوي للتطبيقات اللاسلكية التي تحتوي على مصادر طاقة محدودة. كبسولة التنظير اللاسلكية هي أداة طبية أدخلت عام 2000 تستخدم من قبل الأطباء لتشخيص أمراض الجهاز الهضمي. هذه الأداة لا يمكن أن تصور الجهاز الهضمي كله لأن عمر البطارية محدودة. واحدة من الحلول المقترحة هي استخدام نظام ضغط بسيط و فعال للحد من عدد البتات التي تنتقل عن طريق الهوائي و التي تستهلك معظم الطاقة من البطارية. في هذه الأطروحة، نقتراح أنظمة ترميز الفيديو الموزع جديدة بسيطة وفعالة و قابلة للتطبيق للصور من كبسولة التنظير. و أنظمة DVC المقترحة هي أداة ترميز منخفضة التعقيد لأنها تستخدم نوع واحد من الصور بالمقارنة مع أنظمة DVC التقليدية التي تستخدم نوعين من الصور. و تستخدم كل من أنظمة DVC المقترحة قاعدة البيانات لتوليد المعلومات الجانبية (SI) التي تستخدم لفك ترميز صور المصدر. الاختيار المحرز في قاعدة البيانات باستخدام معاملات DC في المساهمة الأولى و الصورة الجزئية في الثانية. و بالنسبة للمساهمة الثالثة، يستخدم برنامج الترميز الكميات المتجهة لتوليد المعلومات الجانبية (SI). يتم تنفيذ عدة خوارزميات في الأنظمة المقترحة لتحسين كفاءة الضغط. النتائج التي تم الحصول عليها هي كبيرة و تثبت الأفكار المقترحة من خلال مقارنتها مع الترميز JPEG المستخدم بالفعل في كبسولة التنظير.

الكلمات البحث: توزيع الفيديو الموزعة، كبسولة التنظير، قاعدة البيانات، ناقلات تكميم، معلومات جانبية.

Table des matières

Liste des tableaux et des figures	VII
Liste des abréviations	XI
Introduction générale	1
1 Codage vidéo distribué : l'état de l'art	
1.1 Introduction.....	06
1.2 La théorie d'information.....	07
1.2.1 L'entropie.....	07
1.2.2 Codage du canal.....	09
1.3 Codage source distribué.....	11
1.3.1 Codage source distribué sans perte.....	11
1.3.2 Codage de source à perte avec l'information adjacente.....	14
1.4 Le codage vidéo traditionnel.....	16
1.5 Le codage vidéo distribué.....	18
1.5.1 Les architectures proposées dans le codage vidéo distribué.....	19
1.5.1.1 Architecture de PRISM.....	19
1.5.1.2 Architecture de Stanford.....	22
1.5.1.3 Comparaison entre les premières architectures WZ.....	25
1.5.1.4 Amélioration du système Stanford : « DISCOVER ».....	26
1.5.1.5 Architecture hybride.....	29
1.6 Évolution du codage vidéo distribué.....	30
1.7 Les applications du codage vidéo distribué.....	32
1.8 Conclusion.....	34

2 La capsule endoscopique

2.1	Introduction.....	35
2.2	Histoire de la capsule endoscopique.....	35
2.3	Les applications.....	38
2.3.1	Capsule de l'intestin grêle.....	39
2.3.2	Capsule œsophagienne.....	41
2.3.3	Capsule colique.....	42
2.4	Technologie.....	42
2.4.1	La capsule.....	43
2.4.2	Le dispositif d'enregistrement.....	47
2.4.3	La station de travail.....	48
2.5	Les défis scientifiques.....	49
2.6	Conclusion.....	52

3 Codage vidéo distribué basé sur la base de données

3.1	Introduction.....	53
3.2	Similarité entre les images de la capsule endoscopique.....	54
3.3	Utilisation le coefficient DC comme une information de hachage.....	56
3.3.1	La construction de la base de données.....	58
3.3.2	L'encodeur.....	59
3.3.2.1	Transformation cosinus discret.....	60
3.3.2.2	La quantification.....	61
3.3.2.3	Codage Golomb-Rice.....	62
3.3.2.4	Codage Slepian-Wolf.....	62
3.3.3	Décodeur.....	64
3.3.3.1	Génération de l'information adjacente.....	64
3.3.3.2	Décodage Slepian-Wolf.....	65
3.3.3.3	Reconstruction.....	66
3.4	Utilisation d'image sous-échantillonnée comme une information de hachage.....	67

3.4.1	L'encodeur.....	67
3.4.2	Décodeur.....	69
3.5	Résultats expérimentaux.....	72
3.5.1	La qualité de l'information adjacente.....	72
3.5.2	La performance de codage du système.....	78
3.6	Conclusion.....	85
4	La création de l'information adjacente utilisant la quantification vectorielle	
4.1	Introduction.....	86
4.2	La quantification vectorielle.....	87
4.2.1	Définition de la quantification vectorielle.....	87
4.2.2	Conception classique de la quantification vectorielle.....	87
4.3	L'architecture du DVC proposée.....	88
4.3.1	L'encodeur.....	88
4.3.2	Décodeur.....	90
4.3.2.1	Création du dictionnaire.....	93
4.3.2.2	Génération de l'information adjacente.....	93
4.3.2.3	Modélisation de corrélation.....	95
4.3.2.4	Décodage Slepian-Wolf.....	95
4.4	Analyse de la complexité de l'encodeur	98
4.5	Résultats expérimentaux.....	103
4.5.1	La qualité de l'information adjacente.....	104
4.5.2	La performance de codage du système.....	109
4.5.3	Test de contrôle d'énergie.....	112
4.6	Conclusion.....	115
	Conclusion générale	116
	Bibliographie	119

Table des figures

Figure 1.1	Un schéma général d'un système de communication	10
Figure 1.2	Un bloc du diagramme schématique du problème SW du codage séparé des sources corrélées.	12
Figure 1.3	Théorème de Slepian-Wolf.	13
Figure 1.4	Illustration du théorème de Slepian-Wolf pour deux sources.	14
Figure 1.5	Codage de source WZ avec information adjacente.	15
Figure 1.6	Les applications sans fil : la caméra surveillance sans fil, le téléphone portable et la capsule endoscopique.	16
Figure 1.7	L'architecture du codeur PRISM	20
Figure 1.8	Le codage par syndrome et le codage entropique dans un seul bloc de 8×8 .	20
Figure 1.9	L'architecture du décodeur PRISM.	22
Figure 1.10	L'architecture de codage vidéo distribué du Stanford.	23
Figure 1.11	Le codec de DISCOVER.	26
Figure 1.12	Les matrices de la quantification	27
Figure 1.13	Codage et position des blocs 4×4 du WZ (blanc) et clé (noire)	29
Figure 1.14	(Gauche) Caméra sans fil ordinaire. (Droite) webcam sans fil portable imite des caméras de surveillance.	32
Figure 1.15	(Gauche) Écran de vidéoconférence, (Droite) caméra vidéo jetable.	33
Figure 1.16	La capsule endoscopique sans fil (<i>Pillcam SB</i>)	34
Figure 2.1	Le premier prototype de la capsule endoscopique.	37
Figure 2.2	Images du (a) estomac et (b) intestin grêle prend par une capsule endoscopique sans fil.	39
Figure 2.3	(a) capsule l'intestin grêle, (b) capsule œsophage, (c) capsule colique	40
Figure 2.4	Système de capsule endoscopique : capsule, dispositif d'enregistrement et poste de travail.	43
Figure 2.5	Structure de la capsule endoscopique.	44
Figure 2.6	Capteurs et boîte d'enregistrement	48

Figure 2.7	Poste de travail.	50
Figure 2.8	Un prototype d'une capsule robotique.	51
Figure 3.1	Exemple de notre idée	55
Figure 3.2	Architecture de DVC proposé pour la capsule endoscopique	57
Figure 3.3	Conversion de l'image en plans de bits	63
Figure 3.4	Réorganisés les bits de syndrome selon l'ordre de codeur	66
Figure 3.5	Architecture du DVC proposé dans [85]	68
Figure 3.6	Diagramme de codage des coefficients DC.	70
Figure 3.7	Matrices de la quantification pour les composants couleurs.	70
Figure 3.8	Qualité de l'information adjacente de la capsule vidéo test 1 et capsule vidéo test 2 par une BDD contient 4483 images.	73
Figure 3.9	Qualité de l'information adjacente de la capsule vidéo test 2 génère par deux BDD.	74
Figure 3.10	Qualité d'image sélectionnée à partir de la BDD (Coefficient DC vs image sous-échantillonnée)	75
Figure 3.11	Qualité de l'information adjacente génère utilisant l'information du hachage et l'image de la BDD	77
Figure 3.12	Résultats expérimentaux sur les capsules endoscopiques : (a) "Capsule Test Video 1" et (b) "Capsule Test Video 2".	78
Figure 3.13	Qualité visuel à partir de la séquence décodée "Capsule Test Video1" et "Capsule Test Video2": (a) image original (b) image décodé par le codec proposé.	80
Figure 3.14	Résultats expérimentaux d'une capsule endoscopique sans fil : (a) "Capsule Test Vidéo1" et (b) "Capsule Test Vidéo 2" pour le composant luminance.	81
Figure 3.15	Résultats expérimentaux sur d'une capsule endoscopique sans fil : (a) "Capsule Test Vidéo 1" et (b) "Capsule Test Vidéo 2" pour le composant chrominance.	82
Figure 3.16	Résultats expérimentaux sur les capsules endoscopiques : (a) "Capsule Test Vidéo 1" et (b) "Capsule Test Vidéo 2".en codant tous les composants d'image.	84
Figure 3.17	la qualité visuelle d'une image décodée utilise le système DVC [84].	85
Figure 4.1	Conception classique d'un système de la quantification vectorielle.	88
Figure 4.2	la conception de la QV proposé pour le système DVC.	90
Figure 4.3	Schéma du système du DVC proposé.	91
Figure 4.4	Partitions de bloc 16x16.	92
Figure 4.5	L'organisation d'un dictionnaire.	94

Figure 4.6	La qualité de l'information adjacente avec une recherche complète et une recherche arbre (a) "test vidéo 1", (b) "test vidéo 2", (c) "test vidéo 3" and (d) "test vidéo 4".	105
Figure 4.7	La qualité visuelle de l'information adjacente.	108
Figure 4.8	les courbes de la performance RD des séquences (a) "Test vidéo 1", (b) "Test vidéo 2", (c) "Test vidéo 3" and (d) "Test vidéo 4" (256x256 pixels avec 2 images/Sec).	110
Figure 4.9	La performance visuelle avec un PSNR de luminance et l'indice SSIM pour "Test Video1", "Test Video2", "Test Video3" et "Test Video4" : (a) trame de l'information adjacente, (b), (c),(d) et (e) DVC-VQ Trames décodé avec débit différent.	113
Figure 4.10	Exemple de codage d'une région d'intérêt avec une haute qualité et une région non informative avec une basse qualité.	115

Liste des tableaux

Tableau 2.1	Systemes et specifications de capsules sans	41
Tableau 4.1	Les propriétés de la qualité de l'information adjacente de la figure 4.6	106
Tableau 4.2	La qualité moyenne de l'information adjacente	107
Tableau 4.3	La comparaison des débits entre les modèles d'estimation d'erreur utilisant le mode on ligne et hors-ligne.	111

Liste d'abréviations

AC	Alteaternating current
AVC	Advanced Video Coding
BCH	Bose-Chaudhuri-Hocquenghem
BD	Bjøntegaard Delta
BDD	Base de données
bpp	bits per pixel
bps	bits per second
BSC	Binary Symmetric Channel
CCD	Charge Coupled Device
CMOS	CoMplementary Oxide Silicone
CPU	Central Processing Unit
CRC	Cyclic Redundancy Check
dB	Décibel
DC	Direct Current
DCT	Discrete Cosine Transform
DISCOVER	DIStributed COding for Video sERvices
DSC	Distributed Source Coding
DVC	Distributed Source Coding
GI	Gastrointestinal
GPGPU	General-purpose computing on graphics processing units

GPU	Graphics processing unit
ICM	Interpolation et la compensation de mouvement
IDCT	Inverse Discrete Cosine Transform
ITU-T	International Telecommunication Union Telecommunications sector
JPEG	Joint Photographic Experts Group
LDPC	Low-Density Parity-Check
LDPCA	Low-Density Parity-Check Accumulate
LLR	Log-Likelihood Ratio
MJPEG	Motion Joint Photographic Experts Group
MSB	Most Significant Bit-plane
MSE	Mean Squared Error
PRISM	Power efficient Robust hIgh compression Syndrome based Multimedia coding
PSNR	Peak Signal-to-Noise Ratio
QM	Quantization Matrix
QP	Quantization Parameter
RD	Rate-Distortion
SAD	Sum of Absolute Differences
SI	Side Information
SW	Slepian-Wolf
WZ	Wyner-Ziv

Introduction générale

Contexte et Motivation

Au cours des dernières années, les techniques multimédias sans fil (PC-portable, téléphone mobile, Tablette, réseaux de capteurs, Internet of Things...) deviennent de plus en plus indispensables dans notre vie quotidienne. À cet égard, le design d'un système efficace pour la compression de la vidéo, dans ces systèmes, doit contenir quelques exigences comme un encodeur simple, une robustesse contre les erreurs de transmission et de l'évolutivité. Les réponses sur ces exigences peuvent être trouvées dans les résultats de la théorie de l'information, selon laquelle une compression efficace est possible. En plus, les techniques de compression des données audiovisuelles se sont beaucoup améliorées pendant les dernières décennies permettant la banalisation des appareils multimédias grand public (MP3, DVD, Blu-ray, Streaming, appareils photo numériques). Il y a peu de gens aujourd'hui qui n'ont pas entendu les noms de standards de compression comme MP3, MP4, MPEG ou JPEG.

Intéressons-nous au paradigme actuel de compression des données vidéo. En règle générale, la technique de codage se base sur les outils suivants :

- Un estimateur de mouvement pour exploiter la redondance temporelle (entre les images) ;
- Une transformation pour exploiter la redondance spatiale dans l'image (typiquement transformée en cosinus discret « DCT ») ;
- Un quantificateur des coefficients de transformé pour exploiter l'imperfection d'un œil humain ;

- Un codeur entropique pour exploiter la redondance statistique dans les coefficients de quantification.

Notons que les codeurs vidéo actuels comprennent également un mode de prédiction spatiale, c'est-à-dire de prédiction d'un bloc à partir des autres blocs dans la même image. Cette prédiction est faite naturellement dans le codeur, car il dispose de l'image à prédire ainsi que les images servant de référence. Bien qu'en général la prédiction soit l'étape qui permette de gagner plus de compression, c'est aussi l'étape la plus complexe. Par conséquent, dans le paradigme actuel le codeur est beaucoup plus complexe que le décodeur.

Ce déséquilibre convient parfaitement au cas de la liaison descendante. Dans ce type de scénario, l'information est codée par un émetteur aux grandes ressources et peut-être ensuite décodée plusieurs fois (par exemple : la diffusion vidéo, stockage de vidéo, etc.). Cependant, aujourd'hui ce modèle est remis en question par l'utilisation de plus en plus répandue des réseaux sans fil, ainsi que des appareils à consommation faible d'énergie avec des capteurs vidéo intégrés, donc en liaison montante. Ceci suppose un codeur à complexité limitée par les capacités d'un appareil (énergie, mémoire), tandis que le décodeur (souvent installé sur une machine puissante de type serveur) n'a pas de contraintes de complexité (dans la limite du raisonnable).

C'est pourquoi les regards se sont tournés ces dernières années vers un nouveau paradigme de codage appelé le codage distribué. Ce paradigme se base sur les résultats théoriques établis par Slepian-Wolf (SW) [1] en 1973 pour le codage sans perte, puis étendus par Wyner-Ziv (WZ) [2] au cas du codage avec pertes en 1976. Ces résultats montrent qu'en théorie (sous certaines conditions) l'exploitation de la redondance peut être déplacée du codeur au décodeur sans que cela impacte la performance totale du système de codage. Ainsi pour le codage vidéo, l'étape de prédiction avec l'estimation de mouvement n'est plus indispensable au codeur : la redondance temporelle peut être exploitée par le décodeur uniquement.

L'organisation de la thèse

Dans cette thèse, nous proposons des architectures du codage vidéo distribué (*Distributed Video Coding: DVC*) pour coder des trames de la vidéo dans les capsules endoscopiques (CE).

Les capsules endoscopiques (CE) sont des dispositifs, non invasifs, qui permettent d'acquérir des images du tube digestif. Un des problèmes des capsules est leur durée de vie, qui est conditionnée par l'énergie consommée pour la prise des images et leur transmission à l'extérieur. En effet, alors que la durée de transit des capsules, tout le long de l'appareil digestif, est de 24 heures environ la durée de vie de la batterie n'est que de 8-10 heures environ. La solution pour augmenter la durée de vie de la CE est d'assurer une bonne gestion de l'énergie disponible. Un des moyens efficace pour la gestion d'énergie est l'utilisation des algorithmes de traitement et de communication efficaces, car une grande partie de l'énergie de la batterie est consommée par la transmission d'images de la capsule au récepteur [3]. L'énergie de transmission dépend du nombre de bits transmis. Par conséquent, la réduction de la quantité de bits transmis (par compression) réduira l'énergie consommée. Le codage vidéo distribué (DVC) permet de réaliser un codage moins complexe, robuste et avec moins de latence pour ces dispositifs. Dans les systèmes classiques de compression, où le codeur est responsable de l'exploitation des corrélations entre les sources, ce dernier est 5 à 10 fois plus complexe que le décodeur. Donc pour les CE, il est nécessaire d'envisager une autre architecture de codes permettant, à qualité équivalente, de simplifier la partie codage quitte à complexifier la partie décodage. La stratégie de DVC est de compresser chaque source toute seule indépendamment des autres, avec des codeurs aussi simples que possible, et de n'exploiter leurs statistiques qu'au niveau du décodeur. Ainsi, la complexité du codeur est déplacée vers le décodeur. Donc c'est une technique qui convient très bien à la capsule endoscopique, qui contient des ressources de traitement et une batterie très limitées.

Le but de notre travail est de réaliser un système de codage vidéo distribué (DVC) qui contient un codeur simple et applicable pour les images endoscopiques (par capsule). Le système DVC proposé doit être simple et efficace avec un taux de compression élevé.

Contributions principales

Pour une complexité de codage faible, le DVC décale les opérations coûteuses vers le décodeur, mais les architectures DVC souffrent de performances de compression inférieures par rapport aux systèmes de codage vidéo conventionnels. Dans cette thèse, on a proposé un nouveau concept du DVC pour améliorer la capacité de la compression dans la capsule endoscopique.

La première partie du travail concernera la théorie de l'information liée au codage source conventionnel et distribué. Plus précisément, ce chapitre formule théoriquement les thèmes de la compression conventionnelle des sources, du compromis codage-distorsion, le codage Slepian-Wolf (SW) et aussi les problèmes du codage Wyner-Ziv (WZ). Ainsi qu'une description succincte des premiers codecs vidéo WZ développés comme l'architecture du Stanford et l'architecture du PRISM et une revue des développements qui sont venus après les premiers codecs vidéo WZ, comme dans l'architecture DISCOVER. Enfin, on a présenté l'évolution et les applications du système DVC jusqu'à aujourd'hui.

Le deuxième chapitre est dédié à la description générale des capsules endoscopiques, où l'on a présenté les différents types de capsules, leurs compositions, leur mode d'utilisation ainsi que les futures capsules endoscopiques. Aussi, nous avons montré l'importance de la compression efficace dans les capsules endoscopiques pour réduire la consommation d'énergie.

Dans le chapitre trois, deux systèmes de DVC sont proposés pour la capsule endoscopique (CE). D'abord, nous avons expliqué notre idée qui se base sur la similarité des corps humains. Ensuite, nous avons présenté la première contribution du DVC (l'encodeur et le décodeur) où l'on a utilisé les coefficients DC comme une signature des blocs d'image (information de hachage). L'architecture proposée n'utilise pas l'estimation du mouvement pour produire l'information adjacente (*side information* : *SI*) dans le décodeur, mais elle remplace l'estimation du mouvement par une base des données (BDD) qui contient des images de la CE. En outre, nous avons expliqué tous les blocs de codage et de décodage en détail comme la DCT, la quantification, codage SW, la DCT inverse et la reconstruction. Puis, nous avons présenté une autre contribution en utilisant une image sous-échantillonnée comme une information de hachage. Dans cette contribution, on a expliqué comment utiliser l'information de hachage et la BDD pour générer l'information adjacente. Enfin, les résultats sont représentés et discutés en détail.

Une autre architecture du DVC est présentée dans le chapitre quatre. L'architecture proposée utilise la quantification vectorielle pour générer l'information adjacente (SI). Le but de cette architecture est d'améliorer la qualité de SI en utilisant des techniques de raffinement. Dans notre système, le codage scalable est entièrement contrôlé par le décodeur et peut-être combiné avec

des algorithmes complexes d'allocation optimale d'énergie. Dans ce cas, il est possible pour les médecins ou pour un algorithme automatique de décider à tout moment d'arrêter le décodage d'une image particulière, par exemple, lorsque la capsule ne se déplace pas ou se déplace très lentement, une comparaison de l'image de SI ou d'une image partiellement décodée par DVC avec les images précédentes peut déterminer si c'est la même image. Lorsque c'est le cas, les améliorations supplémentaires de l'image sont jugées inutiles et le codage peut être arrêté. D'autre part, un algorithme de détection d'anomalie, un saignement par exemple [4-5], peut imposer le décodage d'images de très haute qualité. Enfin, étant donné que toutes les images de GI ne sont pas visées par les médecins, les images de haute qualité ne peuvent être décodées que dans certaines parties du tractus GI. En plus, on a implémenté des algorithmes pour estimer les paramètres de modèle laplacien des erreurs, en ligne. Enfin, les résultats du DVC sont comparés avec le codage traditionnel.

La conclusion générale résume les principales réalisations de notre thèse et elle identifie certaines orientations pour les travaux futurs dans le domaine de codage WZ.

La production scientifique

- D.E. Boudechiche, S. Benierbah and M. Khamadja, "A new approach to generate side information for distributed video coding in capsule endoscopy", *in 6th Int. Symposium On Communication, Control and Signal Processing (ISCCSP)*, Athens, Greece, May 2014.
- D.E. Boudechiche, S. Benierbah and M. Khamadja, "Distributed Video Coding of the Chroma Components for the Capsule Endoscopy images", *in Proceedings Conference ICNAS, Annaba Algeria 2015*.
- Boudechiche, DjamelEddine, Said Benierbah, and Mohammed Khamadja. "Distributed video coding based on vector quantization: Application to capsule endoscopy," *Journal of Visual Communication and Image Representation*, Vol.49, pp.14-26, November 2017.

Chapitre 1

Codage vidéo distribué : l'état de l'art

1.1. Introduction

Les nouvelles applications vidéo telles que les caméras vidéo mobiles, les réseaux de la surveillance sans fil, la capsule endoscopique sans fil et les réseaux de capteurs sont des applications qui doivent contenir un codeur simple avec une très faible consommation de l'énergie. Dans ce contexte, le codage vidéo distribué (DVC), qui est une réalisation particulière du codage de source distribuée (*Distributed Source Coding : DSC*), est devenu une approche très prometteuse pour la réalisation de ces codeurs à faible complexité et à consommation réduite.

Dans ce chapitre, on va introduire le codage vidéo par la théorie d'information où l'on a défini les notions des bases de la théorie comme l'entropie, l'entropie conjointe et la théorie de DSC. Ensuite, on va aborder les premières architectures du DVC ainsi que les dernières améliorations qui leur ont été apportées.

Chapitre 1. Codage vidéo distribué : l'état de l'art

1.2. La théorie de l'information

La théorie de l'information est une branche de la théorie de la probabilité avec une vaste application dans les systèmes de communication. Alors que plusieurs branches sont d'origine mathématique, la théorie de l'information a une origine physique. Elle a été initiée par les chercheurs de la communication pour étudier les équipements électriques de la communication. Elle a été principalement fondée par Claude Shannon en 1948 [6]. Dans sa théorie, Shannon a développé les limites essentielles de la compression des données et de la transmission fiable sur le canal bruité. La théorie de codage des données porte principalement sur des méthodes explicites pour la transmission ou le stockage des données d'une façon efficace et fiable, qui peuvent être divisées en techniques de la compression de données et du contrôle d'erreur.

Les mesures d'information réalisées par Shannon concernent l'entropie et l'entropie conditionnelle. Elles sont les mesures les plus importantes de la théorie de l'information. Shannon a considéré aussi le cas d'existence de plusieurs sources (X et Y) qui sont encodées et décodées conjointement, et il a montré que s'il existe deux codeurs pour X et Y, le taux de perte minimum d'information est donnée par l'entropie conjointe $H(X, Y)$. En 1973, Slepian et Wolf (SW) [1] ont établi que les deux sources (X et Y) peuvent être codées séparément sans perte, à condition qu'elles soient décodées conjointement et que chaque codeur connaisse les statistiques de l'autre source. En 1976, Wyner et Ziv (WZ) ont appliqué les résultats des travaux de SW au cas des sources à valeurs continues et une compression avec pertes [2]. Les deux travaux ont marqué le début de la théorie du codage de sources distribuées (DSC).

1.2.1. L'entropie

La quantité $I(x)$ définie par :

$$I(x) = \log_b \left(\frac{1}{p(x)} \right) = -\log_b p(x) \quad (1.1)$$

$p(x)$ est la distribution de l'événement $X = x$, ou l'incertitude associée à cet événement.

L'entropie $H(X)$ d'une variable aléatoire X avec la fonction de probabilité $p(x)$ est définie par :

$$H(x) = \sum_x p(x)I(x) = -\sum_x p(x)\log_2 p(x) \quad (1.2)$$

Chapitre 1. Codage vidéo distribué : l'état de l'art

Donc, l'entropie est une mesure de l'incertitude moyenne de la variable aléatoire. Elle donne le nombre de bits en moyenne nécessaires pour décrire la variable aléatoire.

L'entropie conjointe et conditionnelle

Soit Y une deuxième variable aléatoire qui prend des valeurs y dans un autre alphabet fini avec la probabilité $\Pr(Y = y)$. Notons avec $p(x, y)$ la loi de distribution conjointe de X et Y :

$$p(x, y) = \Pr(X = x, Y = y) \quad (1.3)$$

L'entropie conjointe $H(X, Y)$ de X et Y peut être écrite comme suit :

$$H(X, Y) = -\sum_{x \in X} \sum_{y \in Y} p(x, y) \log p(x, y) \quad (1.4)$$

Qui peut également être exprimée comme suit :

$$H(x, y) = -E \log p(x, y) \quad (1.5)$$

Où :

$$E = \sum_{x \in X} \sum_{y \in Y} p(x, y) \quad (1.6)$$

On peut aussi définir l'entropie conditionnelle $H(X|Y)$ de X sachant Y , comme l'espérance d'un évènement $X = x$ sachant $Y = y$:

$$H(X|Y) = -\sum_{x \in X} \sum_{y \in Y} p(x, y) \log p(x|y) = -E \log p(x|y) \quad (1.7)$$

L'entropie conjointe de X et de Y est la somme de l'entropie de X et de l'entropie de Y conditionnellement à l'observation X . Il est facile de montrer que :

$$H(X, Y) = H(X|Y) + H(Y) = H(Y|X) + H(X). \quad (1.8)$$

L'information mutuelle

L'information mutuelle est une mesure de la quantité d'informations que contient une variable aléatoire sur une autre variable aléatoire, il s'agit de la réduction de l'incertitude sur une variable aléatoire après la connaissance de l'autre. L'information mutuelle entre les variables aléatoires X et Y est définie par :

$$I(X; Y) = \sum_{x, y} p(x, y) \log \frac{p(x, y)}{p(x)p(y)} = E \log \frac{p(X, Y)}{p(X)p(Y)} \quad (1.9)$$

Chapitre 1. Codage vidéo distribué : l'état de l'art

1.2.2. Codage source et canal

Un schéma d'un système de communication comprenant le codage de source et de canal est montré sur la figure.1.1. Le codage est séparé en deux parties : codage source et codage canal. Cependant, cette séparation est justifiée par la théorie de l'information, qui montre que le codage de source suivi par un codage de canal est optimal.

Codage source

Le codage de source est une partie de la théorie d'information qui étudie les problèmes liés à la représentation la plus compacte possible des données. Le codage de source connu aussi par un autre nom : « la compression de données » qui vise à substituer un message aussi court que possible au message émis par la source mais fidèle également à l'information à transmettre. Il existe deux types de compression : une compression sans perte (quand la reconstruction parfaite de la séquence d'entrée est garantie) et une compression avec perte (dans ce cas on parle de distorsion de reconstruction).

Pour la compression avec pertes, il y a aussi une limite fondamentale qui dépend de la distorsion voulue. Cette limite est donc une fonction de la distorsion. Elle s'appelle la fonction débit-distorsion $R(D)$, et définit le débit minimum théorique (e.g. en bits par symbole) nécessaire pour coder une séquence de données produite par la source en question, pour que cette séquence soit reconstruite avec une distorsion inférieure ou égale à $R(D)$.

La compression sans pertes est le codage d'une séquence de symboles qui permet un décodage parfait. Il peut se voir comme un cas particulier de la compression avec pertes avec distorsion nulle. Cela implique que $R(0) > H(X)$ (pour une source d'alphabet fini).

Codage du canal

Le codage canal a longtemps été reconnu comme un élément important pour la transmission ou le stockage de l'information numérique, et pour lutter contre le bruit causé par des composants électroniques. Les données sont toujours utilisées sous forme codée pour leur transmission par câbles ou sans fil, ou pour leur stockage magnétiques ou optiques, pour assurer un niveau de fiabilité souhaitée des systèmes de communication.

La théorie du codage de canal montre que si on veut transmettre les données dans un canal bruité avec une probabilité d'erreur faible, on doit diminuer le débit de transmission en ajoutant de la

Chapitre 1. Codage vidéo distribué : l'état de l'art

redondance dans les données. Le codage de canal est alors l'inverse du codage de source, car il ajoute la redondance dans les données.

Le théorème du codage canal

Le théorème du codage de canal [7] dit qu'il est possible de réaliser une communication entre deux appareils sans aucune erreur, à condition que le débit de la transmission soit inférieur ou égal à la capacité du canal.

$$R \leq C \quad (1.10)$$

Où R : le débit de transmission et C : la capacité du canal.

Il existe plusieurs types de codes pour le codage du canal comme : les codes BCH, Reed-Muller (RM), les codes cycliques, les codes de Hamming et Reed Solomon (RS), mais c'est le Turbo code et le LDPC code (*Low Density Parity Check*) qui sont utilisés dans le DSC [8].

Débit-distorsion

La théorie du débit-distorsion permet de définir le débit minimal atteignable sous une contrainte d'une distorsion pour une reconstruction d'une séquence donnée. La reconstruction d'une source codée, sans erreurs, est possible à condition que $R > H$. Ceci dans un cas théorique et optimal, mais il existe des cas où l'entropie $H(X)$ peut être trop grande par rapport au débit R. La question est alors de savoir quel est le niveau de la distorsion qu'on est prêt d'accepter entre l'information

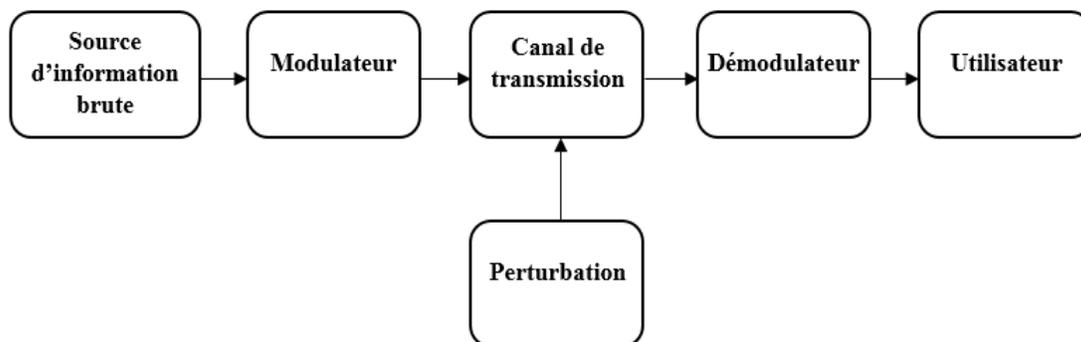


Figure 1.1- Un schéma général d'un système de communication.

Chapitre 1. Codage vidéo distribué : l'état de l'art

brute de la source et l'information en sortie du codeur ? Et quel est le meilleur résultat que l'on peut obtenir lorsque $R < H$? Par conséquent, si $R < H$, quel que soit le code source, la séquence source est presque toujours reconstruite de façon incorrecte.

Alors, la solution est de faire un compromis entre le débit et la distorsion (*Rate-distortion* : $R(D)$), qui permet de reconstruire une source à partir d'un code avec un niveau de distorsion acceptable. Afin de formuler correctement le problème, nous avons besoin d'une mesure de la distorsion entre chaque séquence de la source et chaque séquence reconstruite. Ensuite, nous essaierons de concevoir un code qui a une forte probabilité pour reconstruire une séquence de la source avec un niveau de distorsion toléré.

1.3. Codage source distribué

Le codage source distribué (DSC) est un paradigme du codage dans lequel les sources corrélées sont codées séparément et décodées conjointement. Ce type de codage prend son importance lorsqu'on a des systèmes avec des capacités énergétiques limitées. Le DSC est basé sur deux théorèmes : Slepian-Wolf (SW) et Wyner-Ziv(WZ) [1,2]. Le théorème SW traite le cas où l'on a deux sources statistiquement dépendantes, qui sont codées indépendamment et pas conjointement dans le codeur. Ce théorème affirme que le débit minimal pour coder les deux sources est le même que le débit minimum pour l'encodage conjoint, avec une petite probabilité d'erreur. Donc, le théorème de ouvre les portes à un nouveau paradigme de codage, du moins en théorie. Le théorème WZ étudie le cas où le codeur permet une petite perte d'information dans le codage pour obtenir une compression très efficace.

1.3.1. Codage source distribué sans perte

On considère une compression de deux sources corrélées et discrètes aléatoires X et Y . Dans le codage traditionnel, comme le codage prédictif, l'encodeur et le décodeur exploitent la dépendance statistique entre les deux sources X et Y . Le codage source dit que le débit minimum d'une compression sans perte d'un tel codage est donné par une entropie conjointe $H(Y, X)$. Dans le codage source distribué, l'encodeur code les deux sources X et Y séparément et le décodeur les décode conjointement, exploitant les corrélations entre les sources.

Chapitre 1. Codage vidéo distribué : l'état de l'art

Théorème de Slepian-Wolf

Slepian et Wolf (SW) [1] ont étudié le problème du codage séparé des sources corrélées, qui est représenté dans la figure 1.2. Alors, l'objectif de Slepian-Wolf est de réaliser une compression optimale entre deux sources corrélées sachant que ces deux sources ne communiquent pas durant le codage. On sait que pour une compression optimale sans perte, il faut que le débit de la source X soit supérieur à son entropie: $R_x > H(x)$, mais DSC essaye de réaliser une transmission fiable avec un débit $R_x < H(x)$. La question essentielle est de savoir comment exploiter la redondance entre les deux sources X et Y , si les deux sources ne communiquent pas dans l'encodeur? La réponse à cette question selon SW est que l'encodage de X nécessite un débit R_x tel que :

$$H(X|Y) \leq R_x \leq H(X) \quad (1.11)$$

La figure 1.3 explique la réponse de Slepian-Wolf (SW).

Théorème (SW) : on a deux sources discrètes X et Y défini sur \mathbb{R} :

$$\begin{cases} R_x + R_y \geq H(X, Y) \\ R_x \geq H(X|Y) \\ R_y \geq H(Y|X) \end{cases} \quad (1.12)$$

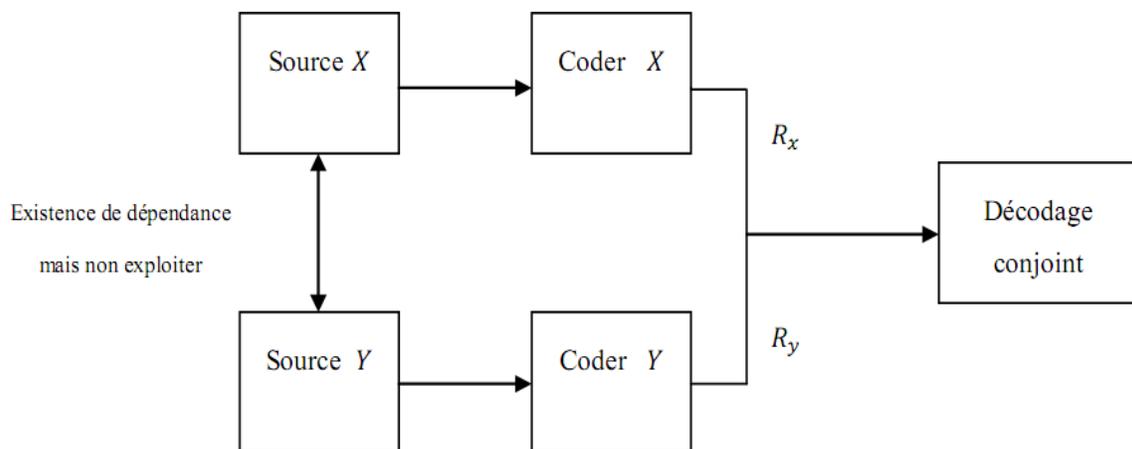


Figure 1.2.-Un bloc du diagramme schématisant le problème SW du codage séparé des sources corrélées.

Chapitre 1. Codage vidéo distribué : l'état de l'art

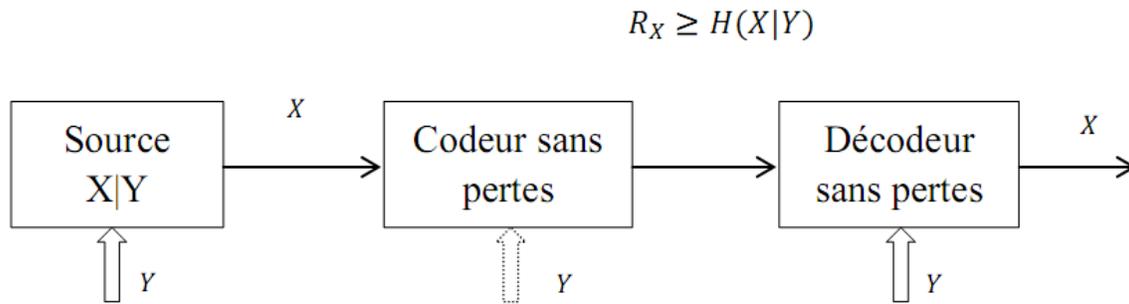


Figure 1.3- Théorème de Slepian-Wolf

Le théorème de SW dit que si on a deux sources à transmettre sans perte d'information, il faut que les trois conditions (équation 1.12) du SW soient vérifiées. La figure 1.4 donne une illustration du théorème de SW dans le cas de deux sources discrètes.

L'objectif du design d'un codage SW est d'atteindre les limites de compression de (1.12). Le but des premières tentatives dans le codage SW est d'atteindre les points d'angle de la région de débit réalisable (les points A et B dans la figure-1.4). Dans ce scénario asymétrique, une source est comprimée et utilisée comme une information adjacente au décodeur pour décoder une autre source. Pour le scénario symétrique, on compresse la source pour l'utiliser comme une information adjacente entre les points A et B par exemple le point C dans la figure 1.4. Ce cas a une importance pratique et s'appelle le codage distribué avec l'information adjacente (*side information* : SI). La source Y est une information adjacente (SI), qui est utilisée pour décoder la source X . D'après la figure 1.4, on peut transmettre la source X avec un débit R_x minimum de l'entropie $H(X|Y)$, en exploitant la source Y , codée par un débit R_y supérieure à l'entropie $H(Y)$. Plus la corrélation entre X et Y est élevée, plus le taux de compression de X est important. La plupart des systèmes réels de codage sont des systèmes de codage avec SI.

Le codage SW est généralement réalisé en utilisant le codage canal par deux méthodes : la première méthode est basée sur un schéma de syndromes (développée par Wyner [9]) et la deuxième est basé sur un schéma de parité. En pratique, ces codes SW utilisent les codes canal les plus efficaces tel que : le Turbo code [10,11] et le LDPC code [12, 13, 14]. Les codes SW basés sur ces deux codes canal ont montré d'excellentes performances et une compression très proche de la limite théorique du SW et une forte robustesse contre les erreurs de transmission. En

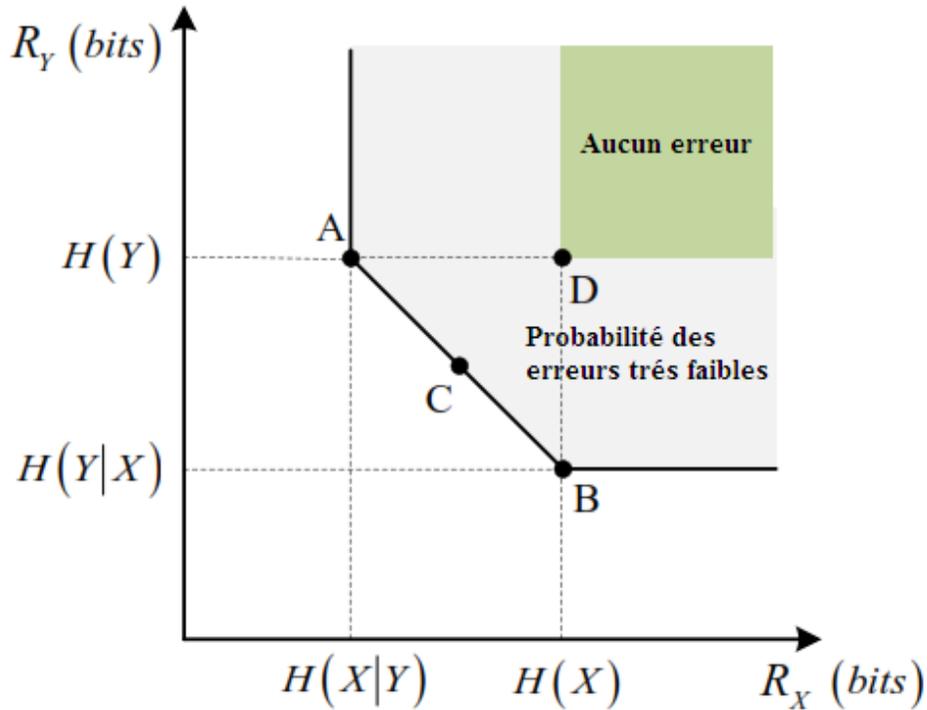


Figure 1.4- Illustration du théorème de Slepian-Wolf pour deux sources.

plus, pour une compression distribuée dans un scénario de transmission sans bruit, le schéma SW basé sur le syndrome avec LDPC est le plus performant [12,13], car il peut atteindre la limite théorique avec une longueur de mot de code plus courte. Par contre, si la transmission est bruitée, le schéma SW basé sur la parité (turbo codes) est préférable pour la compression distribuée, mais il est très complexe que celui basé sur la LDPC [14].

1.3.2. Codage source à perte avec l'information adjacente

Le scénario de compression proposé par SW a stimulé Wyner et Ziv (WZ) à étudier un cas de codage asymétrique du SW où la source X est codée avec perte et avec SI. Donc, le théorème de WZ [2], publiée en 1976, est une extension de la théorie du Slepian-Wolf (SW), mais pour un codage avec perte.

Chapitre 1. Codage vidéo distribué : l'état de l'art

La théorie de Wyner et Ziv

Le théorème de WZ dit que si on a une information de la source X et une information adjacente Y, on peut coder indépendamment la source avec une perte d'information, tel que présentée dans la figure 1.5. Wyner et Ziv ont démontré que :

$$R_X^{WZ}(D) \geq R_{X|Y}(D) \tag{1.13}$$

$R_{X|Y}(D)$: Le débit du codage de la source X si l'information adjacente Y est disponible à la fois à l'encodeur et au décodeur.

$R_X^{WZ}(D)$: Le débit d'encodage de la source X si l'information adjacente est disponible uniquement au décodeur.

Les théorèmes WZ et SW ouvrent de nouvelles possibilités dans le domaine du codage vidéo. L'importance de ces théorèmes est que nous pouvons exploiter la redondance entre deux sources corrélées sans la présence de l'un des deux dans l'encodeur, et ça pour réduire la complexité de l'encodeur. La réduction de la complexité est le facteur le plus important pour les applications sans fil (comme dans la figure 1.6).

En pratique, le codage WZ combine la quantification avec le codage SW pour coder les indices. Dans le codage de WZ, la source a besoin de quantifier X en utilisant un dictionnaire. Le codage

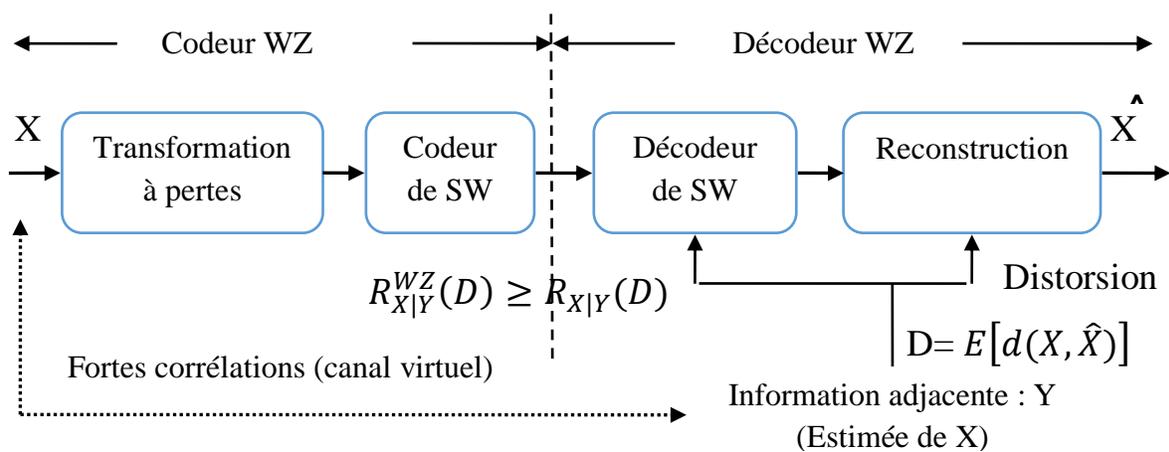


Figure 1.5- Codage de source WZ avec information adjacente

Chapitre 1. Codage vidéo distribué : l'état de l'art

de WZ exploite la corrélation entre la source X et l'information adjacente Y pour obtenir une très bonne compression. Pour atteindre les limites de la théorie du WZ, il est nécessaire d'utiliser deux codes sources, par exemple, la quantification par code de Treillis (TCQ) (qui minimise la perte de codage source en trouvant la meilleure quantification pour un bloc) [15] et les codes de canaux sophistiqués comme le Turbo et LDPC code, qui peuvent s'approcher de limite SW.

Plus précisément, les conceptions initiales du codage WZ, sont axées sur la recherche de bonnes combinaisons entre les codes de Treillis et codage du canal pour le cas gaussien.

1.4. Le codage vidéo traditionnel

Dans les systèmes de codages vidéo traditionnels, de nombreuses méthodes de codage sont utilisées pour exploiter les différents types de redondance dans les signaux tels que : le codage transformé, la compensation de mouvement, la quantification vectorielle, le codage sous-bande...etc.

Parmi les principales réalisations dans ce domaine, on a : le JPEG (1986) pour les images fixes, et Motion JPEG pour la vidéo où les trames sont codées par le JPEG. Les codecs vidéo comme le H.261 (1988), MPEG-1 (1991), MPEG-2 ou H.263, H264 et HEVC qui utilisent deux modes de



Figure 1.6- Les applications sans fil : la caméra surveillance sans fil, le téléphone portable et la capsule endoscopique.

Chapitre 1. Codage vidéo distribué : l'état de l'art

codage (intra et inter) et ils ont établi les fondations des normes actuelles pour le multimédia. Ces normes couvrent une large gamme des services comme les images fixes, la vidéo téléphonique à haute définition et la télévision qui exploitent des débits de quelques dizaines de kbit/s jusqu'à quelques dizaines de Mbit/s.

Par contre, les normes de codage vidéo qui existent aujourd'hui constituent un paradigme où l'encodeur exploite la redondance spatiale et temporelle dans le codage. Ces tâches contiennent des opérations intensives et très complexes en calcul comme l'estimation de mouvement conduisant à des codeurs très complexes. Cette solution correspond à une architecture asymétrique où le codeur est typiquement 5-10 fois plus complexe que le décodeur [16].

Chaque amélioration de ces systèmes réduit le nombre de bits à coder mais rend les systèmes encore plus complexes. Par exemple, le H.264 ou MPEG-4 apporte une grande amélioration estimée à 50 % dans l'efficacité du codage par rapport à la norme H.263. Cette amélioration ajoute une complexité supplémentaire au niveau du codeur, qui utilise de nouveaux outils de codage comme l'estimation et compensation de mouvement avec plus de précision, la prédiction pondérée et de différentes tailles de bloc d'adaptation ainsi que le codage arithmétique. En 2012, une nouvelle normalisation appelée le HEVC (*High Efficient Video Coding*) ou H.265 [17] a apporté une grande amélioration par rapport à H.264 estimée à plus de 50 %. De nouveaux algorithmes ont été ajoutés dans le HEVC comme des techniques de partitions des blocs (33 modes de la prédiction) et d'autres améliorations de précision dans les techniques d'estimation et compensation du mouvement et l'interpolation, qui rend l'encodeur encore plus complexe.

La recherche dans le codage vidéo et la normalisation permet d'ouvrir la porte de recherche scientifique et d'améliorer d'autres paradigmes de compression comme le codage vidéo distribué. En fait, le codage vidéo est l'un des domaines pratiques de la théorie du WZ où l'on peut appliquer les principes du codage source distribué. En effet, on peut exploiter la redondance qui se trouve entre les trames reçues successivement pour la génération de SI. En plus du décodage de canal, l'information adjacente est également utilisée dans le décodeur pour effectuer une reconstruction de la source. De cette façon, l'information adjacente diminue la distorsion de la source reconstruite X' .

1.5. Le codage vidéo distribué

Le but du codage vidéo distribué (DVC) est la conception de systèmes de codage vidéo appropriés qui répondent aux contraintes austères des applications sans fil. Contrairement au codage vidéo traditionnel, dans le DVC, les redondances de la source sont exploitées par le décodeur et non par l'encodeur, ce qui implique un codage simple et un décodage qui peut être complexe. Plus précisément, une prédiction de la source est générée au décodeur en utilisant des informations déjà décodées comme information adjacente. En exprimant la dépendance statistique entre l'information source et l'information adjacente sous forme d'un canal de corrélation virtuel (comme [18] [19]). La compression peut être obtenue en transmettant des bits de parité ou de syndrome d'un code de canal comme le Turbo ou le LDPC code, qui sont utilisés pour décoder la source en utilisant l'information adjacente, qui existe dans le décodeur. Par conséquent, des tâches coûteuses en calcul (comme l'estimation de mouvement) pourraient être déplacées vers le décodeur, permettant un partage flexible de la complexité de calcul entre l'encodeur et le décodeur, permettant ainsi une conception d'architectures de codage vidéo très légères. Les systèmes DVC facilitent la transmission de la vidéo et offrent quelques avantages :

- Le DVC offre un codage simple [18,19], ce qui augmente l'efficacité des architectures de codage et diminue les exigences en puissance du codeur.
- Le codage canal distribué du DVC [20,21] offre une résilience face aux erreurs de canal de communication [22]. Cette fonctionnalité est très attrayante pour obtenir des communications fiables sur des canaux sans fil.
- Le DVC prend en charge l'évolutivité [23]. C'est un avantage clé pour les réseaux hétérogènes de capteurs sans fil, où les ressources et la bande passante fluctuent régulièrement.
- Le DVC facilite un codage efficace pour les systèmes du multi vues, il permet l'exploitation de la corrélation inter visualisation au niveau de décodeur, ce qui décharge le codeur et supprime la communication inter caméra [24, 25].

Chapitre 1. Codage vidéo distribué : l'état de l'art

1.5.1. Les architectures proposées dans le codage vidéo distribué

Les schémas de DVC pratiques codent les images successives de la vidéo séparément, qui sont utilisées par le décodeur pour créer une SI par rapport à l'information de source. Les chercheurs utilisent l'estimation de mouvement dans le décodeur pour générer cette information, en supposant que le décodeur a suffisamment de puissance pour faire le décodage conjoint.

Nous présentons ici, les deux architectures qui furent les premiers systèmes proposés de codage vidéo, à savoir PRISM [19,26] et l'architecture de Stanford [18,27]. Ensuite, nous présenterons en détail le codeur DISCOVER, qui a été à ce jour une des mises en œuvre les plus développées et les plus efficaces pour le DVC [28,29].

1.5.1.1. Architecture de PRISM

Ce système a été proposé par un groupe de chercheurs de l'université de Berkeley en 2002 [19,26], et il est connu sous la dénomination PRISM (*Power-efficient, Robust, hIgh-compression, Syndrome-based Multimedia coding*). L'architecture PRISM est basée sur le codage des blocs d'image, où chaque bloc d'une taille de 8x 8 pixels est codé indépendamment de l'autre. Le système PRISM utilise les différents types de codage de bloc, dans l'encodeur, et utilise aussi l'estimation de mouvement au décodeur.

Dans la première étape de codage, le codeur du système classe les blocs de trame en trois modes : mode « intra », mode « Skip » et mode « Wyner-Ziv ». Pour choisir le mode de codage, le codeur utilise la corrélation entre la trame à coder et la trame précédente pour détecter les blocs qui sont totalement différents pour les coder en mode « intra ». Les blocs corrélés sont codés en mode « Skip » et les blocs qui ne sont pas complètement différents, sont codés en mode « Wyner-Ziv ». Le codeur a besoin d'effectuer plusieurs traitements pour coder les blocs de trame et il a besoin aussi de sauvegarder complètement l'image précédente dans la mémoire. Le système PRISM essaye toujours de garder la complexité à un niveau bas et donne la priorité du codage en mode « Wyner-Ziv ». Ci-dessous, nous expliquons les détails du système de codage et de décodage pas à pas.

L'encodeur :

Les étapes de codage dans le système PRISM, présentées dans la figure 1.7, sont :

Chapitre 1. Codage vidéo distribué : l'état de l'art

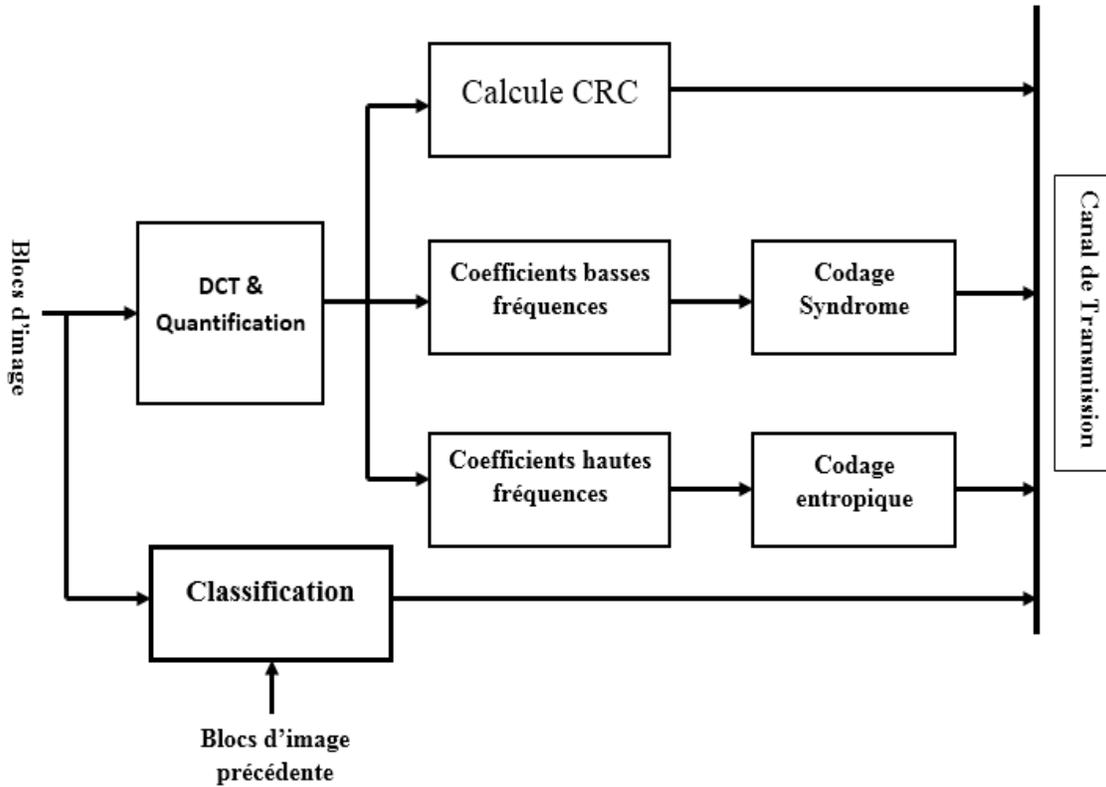


Figure 1.7- L'architecture du codeurPRISM

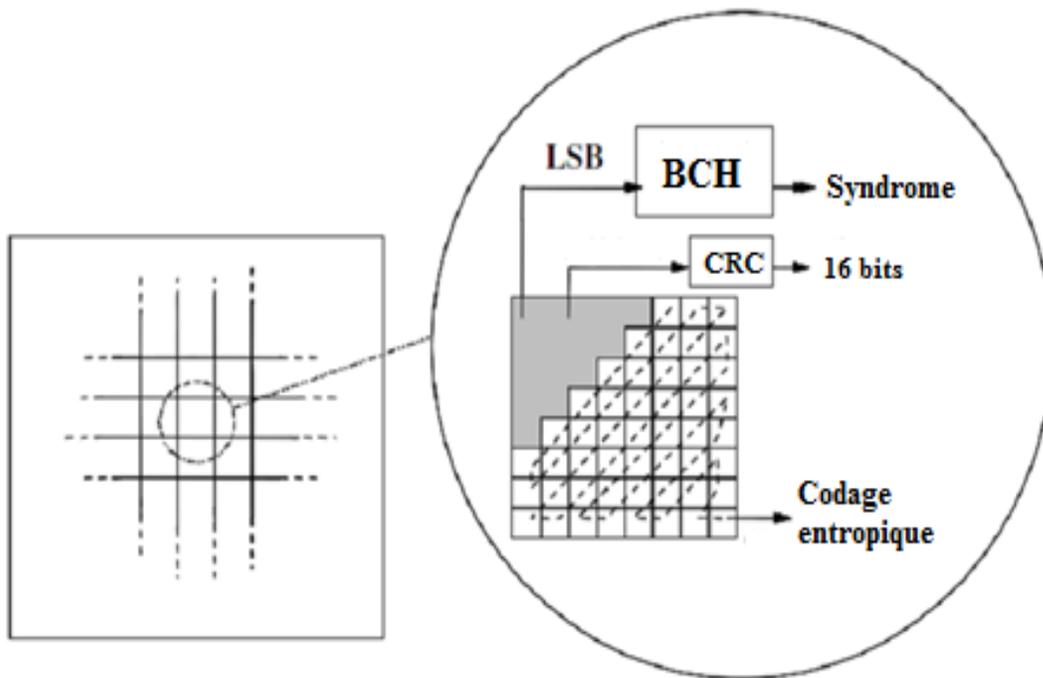


Figure 1.8- Le codage par syndrome et le codage entropique dans un seul bloc de 8x8.

Chapitre 1. Codage vidéo distribué : l'état de l'art

- 1- Chaque trame vidéo est divisée en bloc de taille 8 x 8 pixels, et ensuite la DCT est appliquée sur chaque bloc.
- 2- Un quantificateur scalaire est appliqué aux coefficients DCT correspondants à une certaine qualité ciblée.
- 3- Avant l'encodage, chaque bloc est classé dans une ou plusieurs classes prédéfinies en fonction de la corrélation entre le bloc actuel et le bloc de référence, en fonction de la corrélation temporelle estimée. La phase de classification détermine le mode de codage pour chaque bloc de la trame actuelle : pas de codage (mode « Skip »), le codage intra (mode « intra ») ou le codage du syndrome (mode WZ).
- 4- Les blocs classés dans les classes de codage du syndrome sont codés en utilisant une approche de codage WZ comme décrit ci-dessous. Les modes de codage sont ensuite transmis au décodeur comme une information d'en-tête.
- 5- Pour les blocs qui tombent dans la classe du mode Wyner-Ziv, seuls les coefficients de basse fréquence d'un bloc sont codés en utilisant le codage de syndrome (BCH) voir figure 1.8, et le nombre des coefficients de basses fréquences à envoyer au décodeur dépend de la classe de syndrome à laquelle le bloc appartient. Pour les coefficients de hautes fréquences, ils sont encodés à l'aide d'un codage entropique à base du 4-tuple (exécution, profondeur, chemin, dernier) [20].
- 6- En outre, le codeur envoie pour chaque bloc 16 bits du CRC (détecteur des erreurs de transmission) comme une signature des coefficients DCT quantifiés. Ceci est nécessaire pour sélectionner le meilleur bloc au décodeur pour l'utiliser comme une information adjacente (SI).

Décodeur :

Le décodage des blocs se déroule comme nous pouvons voir sur la figure 1.9 :

- 1- La recherche du mouvement de chaque bloc est faite par rapport à la trame (n-1) précédente, au niveau du décodeur.
- 2- Chacun des blocs joue le rôle SI pour le décodage du syndrome de BCH.
- 3- Chaque bloc candidat conduit à un bloc décodé, à partir duquel une signature de hachage est générée. Afin de détecter un décodage réussi, ce dernier est comparé au CRC reçu à

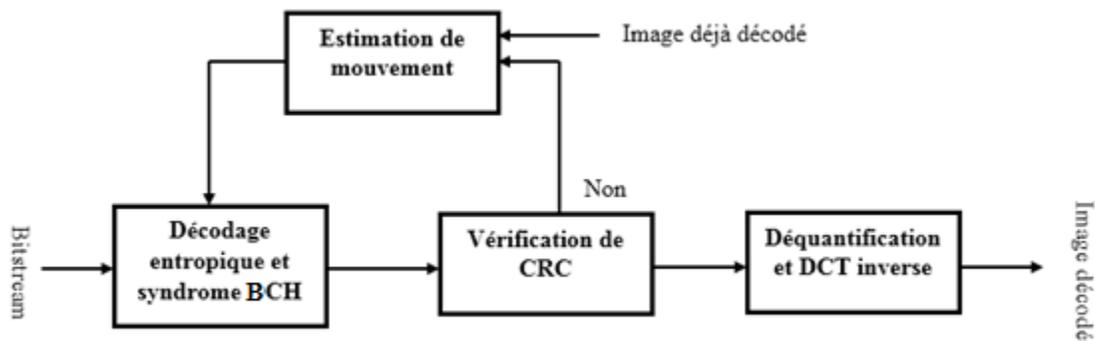


Figure 1.9- L'architecture du décodeur PRISM

partir du codeur. Les blocs de candidats sont visités jusqu'à ce que le décodage mène à une correspondance de hachage.

- 4- Une fois la séquence quantifiée récupérée, elle est utilisée avec les informations adjacentes correspondantes pour obtenir le meilleur bloc reconstruit. Une fois que tous les coefficients transformés ont été déquantifiés, l'opération de balayage en Zig-Zag réalisée au niveau du codeur est inversée pour obtenir un bloc 2-D. Puis, les coefficients transformés sont inversés en utilisant la DCT inverse pour donner des pixels reconstruits.

1.5.1.2. Architecture de Stanford

L'architecture Stanford pour le codage vidéo WZ est la première architecture proposée en 2002 pour le domaine pixels [18] et puis elle est proposée pour le domaine transformé [27]. Cette architecture présente une complexité simple en comparaison à des codecs traditionnels. La figure 1.10 présente l'architecture du Stanford dans le domaine transformé :

L'encodeur :

Classification des images : dans la première étape du codage, la séquence vidéo est divisée entre les trames WZ et les trames clés. Les trames clés sont codées périodiquement, et déterminent la taille du groupe des images GOP (*Group Of Picture*). Les trames clés sont codées par un codeur intra qui exploite seulement la redondance spatiale et sans utiliser l'estimation du mouvement. Par exemple, avec des codeurs de standards H.263+ ou H.264/AVC dans le mode

Chapitre 1. Codage vidéo distribué : l'état de l'art

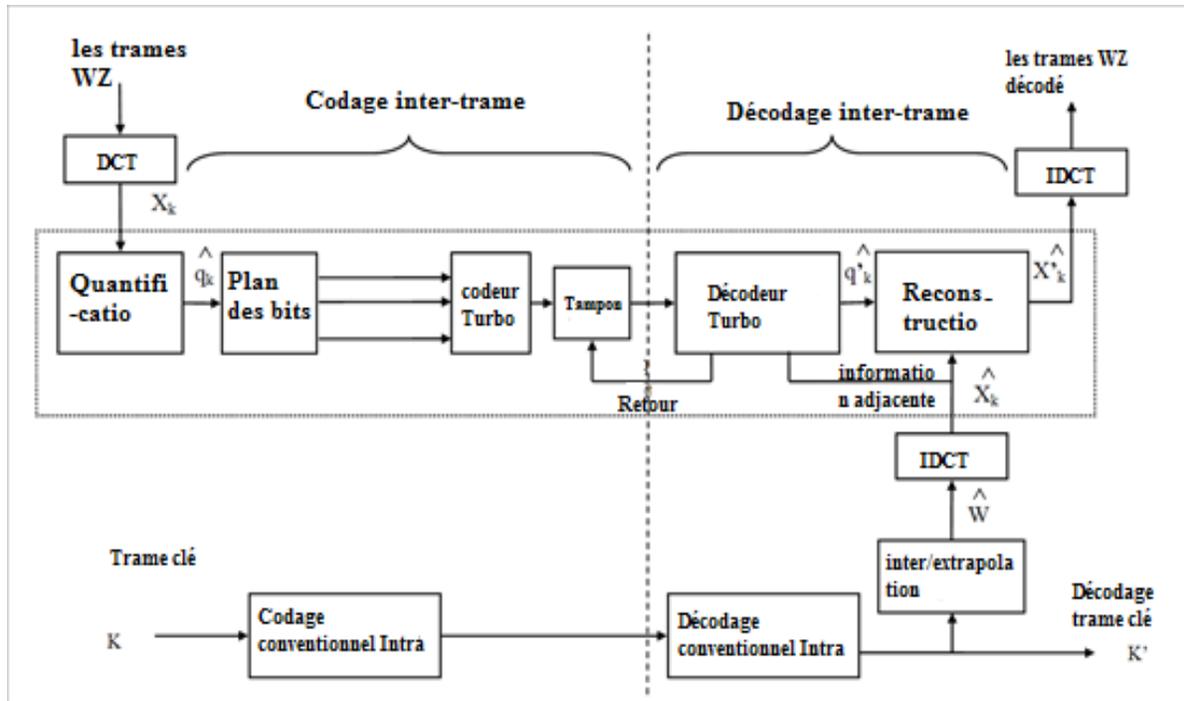


Figure 1.10- L'architecture de codage vidéo distribué du Stanford.

intra. Les autres trames du groupe sont codées par le codeur WZ. Alors, si on a un GOP=2 et 50 trames dans une séquence, le nombre de trames clés est 26 trames et 24 trames WZ.

La transformée DCT : chaque bloc d'une taille 4x4 pixels de la trame Wyner-Ziv subit d'abord une transformation DCT. Ensuite, les coefficients DCT de la trame sont regroupés ensemble et organisés par leurs fréquences (DC, AC1,..., AC15).

La quantification : chaque bande de la fréquence des 16 bandes subit une quantification avec un quantificateur scalaire uniforme à 2^{Mk} niveaux, $k = 1 \dots 16$, où les valeurs sont choisies en fonction de la distorsion. Enfin, les valeurs quantifiées des bandes DCT sont regroupées en plan de bits et encodé par un codeur du canal.

Codeur du canal : le codeur du canal utilisé par ce système c'est le Turbo code. Il commence le codage des plans de bits à partir du plan le plus significatif (*MSB*). Les bits de parité sont générés pour chaque plan de bits, et puis ils sont mis dans un tampon et transmis selon les requêtes du décodeur.

Chapitre 1. Codage vidéo distribué : l'état de l'art

Décodeur :

Création de l'information adjacente : le décodeur crée l'information adjacente (SI) pour chaque trame WZ codée. Il utilise l'estimation et la compensation du mouvement entre les trames clé les plus proches. La SI est considérée comme une estimation (version bruitée) de la trame WZ originale. Une bonne estimation permet de générer une meilleure SI, qui nécessite moins de bits pour la corriger en utilisant le Turbo code.

Estimation des probabilités de canal : les statistiques de la différence (résidus) entre les coefficients de la trame WZ et l'information adjacente sont modélisées par une distribution Laplacienne dont les paramètres ont été estimés précédemment. Dans le système de Stanford, les paramètres de modèle Laplacien sont estimés hors-ligne et le paramètre α du modèle Laplacien est défini comme :

$$\alpha = \sqrt{\frac{2}{\sigma^2}} \quad (1.14)$$

Où σ^2 est la variance du résidu entre la trame WZ et la trame de l'information adjacente.

Décodage du canal : lorsque les coefficients DCT de SI et les statistiques sur les résidus sont connues, le système de décodage commencera à décoder chaque plan de bits par le Turbo code (à partir du MSB). Le turbo décodeur reçoit à partir du codeur des groupes successifs de bits. Pour décider si le décodage réussit ou non, le décodeur utilise un critère d'arrêt. Lorsque le décodeur réussit à décoder le plan de bits MSB d'une bande DCT, le décodeur passe, d'une manière similaire, vers le plan de bits suivant. Une fois que tous les plans de bits d'une bande DCT sont décodés avec succès par le Turbo code, le décodeur commence à décoder la bande suivante.

La reconstruction : après le décodage de tous les plans de bits associés à chaque bande DCT, les plans binaires sont regroupés pour former les flux de symboles quantifiés décodés associés à chaque bande. Une fois que tous les symboles quantifiés décodés sont obtenus, il est possible de reconstituer les coefficients DCT. Enfin, les coefficients de la DCT des bandes non décodées sont remplacés par les bandes DCT correspondantes de SI.

Transformation inverse de la DCT : une fois que toutes les bandes DCT sont reconstituées, la transformation inverse est appliquée (IDCT), et la trame WZ décodée est obtenue.

Chapitre 1. Codage vidéo distribué : l'état de l'art

Assemblage des images : enfin, pour obtenir la séquence vidéo décodée, les trames clés décodées et les trames WZ sont mélangées.

1.5.1.3. Comparaison entre les premières architectures WZ

Les raisons pour lesquelles la communauté de la recherche a adopté avec plus d'enthousiasme l'architecture de Stanford ne sont pas entièrement claires. Il était très probable qu'un facteur pertinent soit : une plus grande littérature est disponible ou une mise en œuvre initiale plus facile. Du point de vue technique, les principales différences fonctionnelles suivantes peuvent être mises en évidence :

- 1- Le codage basé sur la trame (Stanford) et le codage basé sur le bloc (PRISM) : Pour l'architecture de Stanford, les trames de séquences vidéo sont codées complètement en mode Wyner-Ziv ou en mode intra. Par contre dans l'architecture de PRISM, les blocs de trames sont codés en mode Wyner-Ziv, mode Intra ou mode Skip.
- 2- Contrôle du débit du décodeur (Stanford) par rapport au contrôle du débit du codeur (PRISM). Dans le premier cas, le canal de retour est nécessaire pour décoder les bits, ce qui limite la portée aux applications en temps réel.
- 3- L'encodeur très simple (Stanford) contre un encodeur intelligent (PRISM). L'activation des opérations inter-images est limitée à l'encodeur qui permet d'intégrer les décisions en mode de codage variant dans l'espace.
- 4- Les codes de canal comme le Turbo code et le LDPC sont plus sophistiqués (Stanford) par rapport au code de canal plus simple (PRISM) comme le code BCH.
- 5- Aucune donnée auxiliaire, en plus des codes de hachage, n'est envoyée par le codeur pour aider le décodeur dans le processus d'estimation de mouvement dans l'architecture Stanford.
- 6- Moins intrinsèquement robuste à la corruption d'erreur dans Stanford contre une résilience plus élevée à la corruption d'erreur dans PRISM en raison de la recherche de mouvement comme approche effectuée au décodeur.

Avec le temps, certaines différences citées ci-dessus ont disparu, par exemple, il existe actuellement des solutions de codage basées sur l'architecture du Stanford qui utilise les modes des blocs dans la trame WZ [30], et des signatures de hachage (CRC par exemple) transmises par

Chapitre 1. Codage vidéo distribué : l'état de l'art

Le codec DISCOVER utilise l'architecture Stanford dans le domaine transformé [26]. Le groupe DISCOVER a essayé d'améliorer les modules de l'architecture de Stanford comme la transformée de DCT, la quantification, le codage et le décodage de Slepian-Wolf, la modélisation de corrélation et la reconstruction. Ces nouvelles améliorations sont mises en relief dans la figure 1.11. Dans ce codeur, le groupe DISCOVER a ajouté d'autres matrices de quantification, à savoir huit matrices de quantification présentées dans la figure 1.12, afin d'atteindre une plage de débit plus élevée [37]. En outre, un quantificateur uniforme a été utilisé pour les coefficients DC, mais une double zone morte a été utilisée pour les coefficients AC, ce qui entraîne des performances de codage légèrement meilleures [36]. Pour le codage de Slepian-Wolf, au lieu d'utiliser le Turbo code comme dans l'architecture de Stanford, le codec DISCOVER a utilisé le LDPCA (*Low-Density-Parity-Code-Accumulate* : *LPDCA*) de [12]. DISCOVER a ajouté aussi de nouvelles techniques pour augmenter les performances RD dans le système :

- **Technique de contrôle adaptatif pour les groupes des images** : cet algorithme permet d'optimiser la classification des images entrant vers le codeur et classifier en deux types de trames Wyner-Ziv et Clé. Le moment où la vidéo présente un mouvement rapide, le nombre d'un groupe d'images (GOP) est réduit et si le mouvement de vidéo devient lent le groupe d'images est augmenté.

16	8	0	0
8	0	0	0
0	0	0	0
0	0	0	0
M1			
32	8	0	0
8	0	0	0
0	0	0	0
0	0	0	0
M2			
32	8	4	0
8	4	0	0
4	0	0	0
0	0	0	0
M3			
32	16	8	4
16	8	4	0
8	4	0	0
4	0	0	0
M4			
32	16	8	4
16	8	4	4
8	4	4	0
4	4	0	0
M5			
64	16	8	8
16	8	8	4
8	8	4	4
8	4	4	0
M6			
64	32	16	8
32	16	8	4
16	8	4	4
8	4	4	0
M7			
128	64	32	16
64	32	16	8
32	16	8	4
16	8	4	0
M8			

Figure 1.12- Les matrices de la quantification.

Chapitre 1. Codage vidéo distribué : l'état de l'art

- **Algorithme d'interpolation adaptatif** : face à la variation de taille du GOP, l'algorithme de l'interpolation entre les images doit être aussi flexible pour la génération de SI, même si la taille du GOP est élevée.
- **La modélisation en ligne des paramètres du canal virtuel** : le codec DISCOVER permet l'estimation du paramètre α dans le décodeur en ligne pour éviter la génération de SI au niveau de codeur. De plus, DISCOVER utilise l'estimation du paramètre α au niveau de chaque coefficient de la DCT et pas seulement au niveau de trame.
- **Estimation du débit minimal au niveau de l'encodeur** : Une autre technique est proposée dans le codec DISCOVER qui permet de réduire le nombre de paquets demandés en utilisant le canal de retour [38]. Un débit minimal pour décoder la trame WZ est estimé par l'encodeur en utilisant des trames précédentes pour créer une simple SI, de moindre qualité. La génération de la SI dans l'encodeur est faite par un simple algorithme d'estimation de mouvement pour garder l'encodeur moins complexe. Si le décodeur ne peut pas décoder le plan de bits en utilisant la quantité de bits transmis par le codeur SW, il demande des bits complémentaires, avec le canal de retour. De cette manière, le nombre total de demandes de canal de retour du décodeur au codeur est réduit ; diminuant ainsi la latence associée.
- **Amélioration d'estimation du mouvement** : dans le décodeur, le codec DISCOVER dispose d'un outil de génération de SI par rapport à celui développé par le groupe Stanford. Comme dans ce dernier cas, la méthode d'estimation de mouvement crée une prédiction de la trame WZ codée, en utilisant une image précédente et une image future dans une structure de prédiction bidirectionnelle hiérarchique. Cependant, la méthode d'estimation de mouvement de DISCOVER englobe des extensions avancées comprenant une estimation et une compensation de mouvement bidirectionnel basées sur des blocs [39, 40], un raffinement de mouvement de sous-pixel [41] et un lissage spatial des vecteurs de mouvement [38, 40].
- **La reconstruction** : Après le décodage SW, DISCOVER effectue une reconstruction optimale pour obtenir des valeurs des coefficients décodés. Dans ce contexte, une contribution supplémentaire du groupe DISCOVER a été proposée dans l'article [42] pour calculer les valeurs de la reconstruction à l'aide du modèle de la corrélation Laplacien.

Chapitre 1. Codage vidéo distribué : l'état de l'art

1.5.1.5. Architecture hybride

L'architecture hybride proposée par les auteurs de [43] présente un nouveau mode de trame. Alors que les architectures traditionnelles du DVC utilisent deux modes de trames : la trame WZ et la trame Clé, [43] propose de généraliser un codage hybride pour inclure les trames WZ et les trames Clés en tant que mode spécial. D'autres améliorations sont également proposées dans cette architecture, telles qu'une nouvelle quantification efficace et une vérification en utilisant la redondance cyclique (CRC). L'architecture hybride utilise le diagramme du Stanford comme le montre la figure 1.10, le système utilise des macro-blocs de 16x16 et chaque macro-bloc est divisé en des blocs de 4x4 comme le montre la figure 1.13. Le but de ce système est d'améliorer les décisions d'estimation et de compensation du mouvement. De plus, pour chaque macro-bloc (MB), l'estimation du mouvement est possible au décodeur. L'information intrinsèque d'un MB est utilisée pour l'estimation du mouvement avec une correspondance partielle pour trouver une bonne prédiction des parties restantes de même que MB aide à sélectionner les meilleurs modes. Certaines améliorations sont ajoutées dans cette architecture par rapport à d'autres architectures :

- L'estimation du mouvement pour les blocs 8x8 avec une précision de 1/2 pixel ;
- La quantification est modifiée pour faciliter l'utilisation avec un système hybride et multimode. La quantification dans l'architecture DISCOVER est basée seulement sur huit points RD et n'autorise pas la variation du taux, comme dans le codage vidéo conventionnel, et les coefficients de la DCT sont supprimés sans tenir compte de leur amplitude réelle. Dans le système hybride, une seule matrice de quantification de base ($Q_i = 8$) est utilisée. Les valeurs de la matrice de quantification M sont multipliées par un

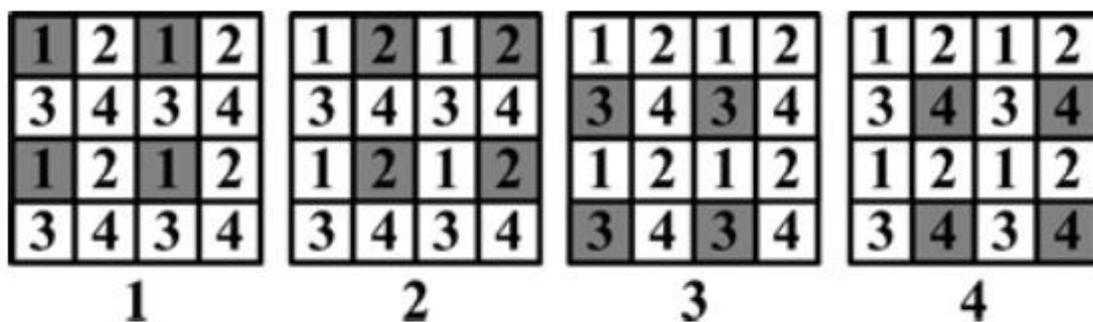


Figure 1.13- Codage et position des blocs 4x4 du WZ (blanc) et clé (noire)

Chapitre 1. Codage vidéo distribué : l'état de l'art

paramètre q_w , résultant en une autre matrice utilisée pour la quantification, le système utilise une seule matrice de quantification. Le changement de la matrice faite par le changement du coefficient q_w pour obtenir de nombreuses qualités, d'une manière similaire à celle du JPEG.

- Pour améliorer le codage, le système utilise le CRC pour vérifier uniquement les huit premiers coefficients de la DCT et pour les plans de bits codés avec un taux de compression supérieur à 2, où la probabilité des erreurs est encore très faible. La plupart du temps, il n'y a pas assez d'erreurs sur le plan de bits et le CRC est perdu. S'il y a une erreur, le CRC est envoyé pour chaque bande. Après avoir déterminé la bande erronée, le CRC est envoyé pour chaque plan de bit pour protéger celui qui est erroné et le corriger.

1.6. Évolution du codage vidéo distribué

La majeure partie de la littérature sur le codage vidéo distribué (DVC) a abordé le problème de la réduction de la complexité du codage, en déplaçant la tâche d'estimation et compensation de mouvement de l'encodeur vers le décodeur. Le problème, par contre, est que les systèmes DVC proposés sont moins efficaces que les systèmes de codage vidéo classiques.

L'efficacité de la compression du DVC dépend fortement de la corrélation entre l'information adjacente et la trame WZ qui est générée par l'interpolation. Dans l'article [44], les auteurs effectuent un lissage spatial avec un raffinement des vecteurs de mouvement pour améliorer la qualité de SI. De même, dans [45], deux SI sont générés par extrapolation des trames clés précédentes et suivantes respectivement, en utilisant des vecteurs de mouvement avant et arrière. Ensuite, le processus de décodage utilise simultanément les deux SI. La précision de sous-pixel est semblable à la méthode de H.264/AVC proposée dans l'article [46] afin d'améliorer davantage l'estimation de mouvement pour la génération de SI. Dans [47], les vecteurs de mouvement sont affinés sur la base du décodage de plan de bits de la trame WZ reconstruite, ainsi que des trames clés précédemment décodées. Il permet également l'utilisation de différents modes d'interpolation. Cependant, seules des améliorations mineures sont signalées. L'approche de [48] partage certaines similitudes du travail de [47]. Une trame WZ partiellement décodée et reconstruite est exploitée pour améliorer itérativement l'interpolation temporelle compensée par le mouvement. Dans l'article [49], les auteurs ont introduit une méthode itérative pour générer SI

Chapitre 1. Codage vidéo distribué : l'état de l'art

avec un raffinement de mouvement. D'autres chercheurs, dans [49-50], ont utilisé une information du hachage pour améliorer la recherche des blocs. Les auteurs dans [51] ont proposé un nouveau codec DVC qui utilise deux méthodes d'estimation du mouvement. La première méthode utilise le flot optique sur le niveau des pixels et la deuxième méthode utilise l'estimation du mouvement traditionnel sur le niveau des blocs. Les résultats de ce système sont les meilleurs sur le plan de la performance RD, en comparaison à tous les systèmes proposés dans le DVC. Cependant, le système [51] est très complexe, parce qu'il régénère la SI après chaque plan de bits décodé.

Un autre facteur important ayant un impact sur la performance du DVC est l'estimation du modèle de corrélation entre les trames SI et WZ. Dans certains schémas DVC [18], le modèle Laplacien est calculé hors ligne, sous l'hypothèse irréaliste que les images d'origine sont disponibles au décodeur. Dans [52], une méthode est proposée pour l'estimation en ligne au niveau du décodeur du modèle de corrélation. Une autre technique est proposée, dans l'article [53], qui consiste à calculer les paramètres du modèle de corrélation à l'encodeur par approximation de SI. Une autre technique est proposée dans [54], pour améliorer l'estimation du paramètre de la distribution Laplacienne après chaque bande et plan de bits décodés.

Le module de reconstruction joue également un rôle important dans la détermination de la qualité de la vidéo décodée. Dans l'architecture Stanford [18], le pixel reconstruit utilise l'information adjacente correspondante et les limites de l'intervalle de la quantification. Une autre approche proposée dans [55] tire parti de la distribution statistique moyenne des coefficients de transformation. Dans [56], la valeur reconstruite est ensuite calculée comme espérance du coefficient source compte tenu de l'intervalle de quantification et de la valeur d'information adjacente, montrant une performance améliorée. Un nouvel algorithme est introduit dans [57], il exploite la distribution statistique du bruit de la sortie DVC décodée.

La performance de l'interpolation du mouvement est également analysée théoriquement dans [58,59]. Dans [59], les auteurs montrent que la précision de l'interpolation dépend fortement de la cohérence temporelle du champ de mouvement, ainsi que de la distance entre les trames. Un modèle basé sur un filtrage de Kalman démontre que le DVC avec interpolation de mouvement au décodeur ne peut atteindre la performance du codage prédictif classique. Une méthode pour optimiser la taille du GOP est également proposée. Dans [60], les auteurs implémentent en temps

Chapitre 1. Codage vidéo distribué : l'état de l'art

réel tout le système du DVC, en utilisant les techniques de traitement parallèle GPGPU qui donnent de bons résultats.

1.7. Les applications du codage vidéo distribué

Le paradigme du DVC est principalement utile pour les applications qui ont besoin d'encodeurs de faible complexité et de faible consommation. C'est le cas par exemple des applications émergentes sans fil qui ont besoin d'un encodeur simple, d'une robustesse de canal de la transmission, d'une compression efficace et élevée, et d'un retard très petit. En plus, les domaines potentiels pour utiliser le paradigme du DVC sont cités ci-dessous :

Caméra de surveillance sans fil

Les applications du réseau de surveillance sans fil ont une faible consommation d'énergie et ils sont principalement axés sur la surveillance et sur la sécurité. Les divers composants du réseau de faible consommation sont interconnectés et la communication entre elles s'effectue à travers des protocoles de communication sans fil. Les composants qui fournissent des informations pour le système sont des caméras qui peuvent capturer ou afficher à un ou plusieurs dispositifs. Alors que certaines applications de surveillance sans fil envisagent une seule caméra, d'autres applications envisagent un scénario multivues où une corrélation intervue est à exploiter. En



Figure 1.14- (Gauche) caméra sans fil ordinaire (Droite) webcam sans fil portable imite des caméras de surveillance.

Chapitre 1. Codage vidéo distribué : l'état de l'art

outre, la qualité et l'évolutivité spatiale peuvent être des problèmes pertinents si le décodeur, pour une vue spécifique, utilise l'image décodée à partir d'une autre vue pour fournir une vidéo décodée avec une qualité ou une résolution accrue. Dans ce cas, la qualité ou la résolution supplémentaire sera fournie en fonction de données qui ne sont pas déterministes.

Vidéoconférence avec périphériques mobiles

La vidéoconférence concerne principalement la transmission d'une image synchronisée (vidéo) et d'une parole (audio) entre deux ou plusieurs emplacements physiquement séparés (voir la figure 1.15, à gauche). Parfois, on ne peut avoir une réunion en face à face avec deux personnes ou plus. D'autre fois, une conversation téléphonique ou une conférence téléphonique est adéquate. La vidéoconférence ajoute une autre alternative possible : (I) une conversation vivante est nécessaire ; II) l'information visuelle est un élément important de la conversation ; III) les parties de la conversation ne peuvent pas physiquement se rendre au même endroit ; (IV) la dépense ou l'heure du voyage est un problème.

Capsule endoscopique sans fil

La capsule endoscopique (CE) est un outil électronique utilisé pour diagnostiquer les maladies du tractus gastro-intestinal humain, les techniques d'examen standard permettent de visualiser l'œsophage et l'estomac (la gastroscopie et la coloscopie). Cependant, la partie la plus longue du système digestif humain, à savoir l'intestin grêle, reste principalement inaccessible au sondage



Figure 1.15- (Gauche) écran de vidéoconférence (Droite) caméra vidéo jetable.



Figure 1.16- La capsule endoscopique sans fil (*PillcamSB*)

avec de telles méthodes de diagnostic invasives. Les progrès récents dans l'électronique, la fabrication de microsystèmes, l'imagerie médicale et les communications sans fil ont conduit à la conception d'une capsule endoscopique sans fil. La capsule contient une batterie miniature, une source d'éclairage efficace, une caméra vidéo et un émetteur-récepteur RF. Une fois prise, la pilule enregistre la vidéo de la piste digestive et transmet à un récepteur monté sur le corps du patient. Actuellement, l'objectif majeur des fournisseurs de la capsule endoscopique (par exemple, Given Imaging, Olympus Optical) est d'obtenir une compression vidéo hautement efficace et résistante aux erreurs, à faible consommation d'énergie, ainsi qu'une durée de vie de la capsule longue.

1.8. Conclusion

Le système du codage vidéo distribué est un nouveau paradigme de la compression. Ce système propose un codeur simple, ce qui le rend candidat pour les applications sans fil et à faibles ressources. Les chercheurs améliorent ce système chaque jour et essayent de réduire ces inconvénients et d'augmenter l'efficacité de la compression.

Dans ce chapitre, on a présenté la partie théorique de la thèse, en montrant les bases théoriques du codage vidéo distribué et en expliquant le fonctionnement des premières architectures DVC. Ensuite, on a présenté l'évolution du système et les domaines d'applications possibles.

Chapitre 2

La capsule endoscopique

2.1. Introduction

Les maladies gastro-intestinales (GI), telles que les cancers de l'estomac et du côlon, constituent maintenant les grandes menaces pour la santé humaine. Par exemple, selon les statistiques de 2002 [61], le nombre de cas de cancer de l'estomac pour les femmes en *Algérie* est classé le deuxième parmi tous les cas de cancer. Les différentes maladies dans le tractus gastro-intestinal peuvent être évitées et guéries en cas de découverte précoce de la maladie. Les méthodes de détection traditionnelles, telles que l'endoscopie, l'échographie et la tomodensitométrie, permettent de déterminer les diagnostics des maladies du tractus digestif, mais elles peuvent avoir certains inconvénients tels que la douleur et l'incertitude.

Dans ce chapitre, on se propose de présenter ce qu'est une capsule endoscopique et d'en faire une description générale : leur composition, leur performance et aussi leurs limites, ainsi que les perspectives dans ce domaine.

2.2. Histoire de la capsule endoscopique

Il y a plus de 50 ans que la visualisation directe dans le système gastro-intestinal (GI) a été développée en utilisant l'endoscopie traditionnelle pour le diagnostic des gastro-entérologies. Ce système permet d'utiliser des écrans vidéo à haute définition qui sont utilisés quotidiennement

Chapitre 2. La capsule endoscopique

pour diagnostiquer le système gastro-intestinal supérieur et inférieur ainsi que pour les voies biliaires. Cependant, l'intestin grêle pose un problème réel pour les médecins, parce qu'ils ne peuvent pas le visualiser directement par l'endoscopie traditionnelle. Une entéroscopie poussée permet une visualisation d'environ 50 % de la longueur de l'intestin grêle.

En 1980, un ingénieur en électro optique, Gavriel Iddan et un gastro-entérologue, Eitan Scapa se sont rencontrés pour discuter de leurs champs de travail, pour essayer de trouver une nouvelle solution technologique qui permette de visualiser la totalité de l'intestin grêle. Après 10 ans, Iddan a utilisé la technologie CCD dans les capteurs d'images remplaçant ainsi la fibre optique. Cette technologie CCD permet à Iddan d'utiliser seulement une caméra en la laissant se mouvoir, naturellement, dans le tractus gastro-intestinal. La deuxième étape consistait à remplacer le câble par un transmetteur. L'idée d'équiper la caméra par un transmetteur a été rapidement développée. Ensuite, une question importante relative au temps nécessaire à la caméra pour traverser le petit intestin s'est posée ? La batterie miniature qui existait alors ne permettait que de 10 à 15 minutes de travail, sans tenir compte de l'énergie nécessaire à la transmission et à l'éclairage.

Ces défis énormes ont presque empêché Iddan de continuer le projet, et il a décidé de s'attaquer à ces obstacles un par un. Il a commencé par fabriquer son premier prototype en utilisant une fenêtre optique "axicon", en supposant que cette forme sera continuellement nettoyée pendant le déplacement. Il a ajouté une caméra miniature CCD et une source lumineuse qui a montré expérimentalement que cela produisait des images raisonnables. Certains de ces premiers modèles sont présentés à la figure 2.1.

En 1993, avec le développement de nouveaux capteurs CCD, qui consomment moins d'énergie, une nouvelle conception est proposée qui divise l'appareil en 3 parties : un capteur CCD, un transmetteur-enregistreur et un poste de travail. Cette conception permet d'analyser les images hors ligne et non en temps réel. Un réseau de capteurs est ajouté à la conception pour la localisation de l'appareil. En 1997, Idana change le capteur d'image CCD par un capteur CMOS qui consomme moins d'énergie et augmente le temps de vie de la capsule.



Figure 2.1-Le premier prototype de la capsule endoscopique

La même année, Idan et Scapa ont travaillé avec un autre groupe de Paul Swain, et ses collègues. Le groupe de Paul Swain a travaillé sur la transmission sans fil et sur la capsule pH. En 1994, Paul Swain a soulevé la possibilité d'un endoscope sans fil dans une conférence donnée lors du congrès mondial de gastro-entérologie. Plus tard, en 1996, le groupe Swain a commencé à tester la transmission d'images vidéo à travers le corps humain.

En 1999, les deux groupes ont procédé aux tests expérimentaux dans la clinique de Scapa avec une capsule de taille 33 x 11 mm. Les résultats obtenus ont montré que le temps de vie ne dépasse pas les 2 h. Plus tard, et avec un très bon transmetteur la capsule a transmis des images de bonne qualité avec un temps de vie allant jusqu'à 6 h et en plus elle a atteint le cæcum. En mai 2000, un petit document dans le journal *Nature* [62] a introduit une nouvelle forme d'endoscopie, appelée la capsule endoscopique sans fil. Cette nouvelle invention, développée par la société *Given Imaging*, a révolutionné presque la méthode du diagnostic dans le tube digestif, car ce petit appareil peut visualiser directement tout l'intestin grêle sans douleur ou sédation. La capsule est présentée pour la première fois au Congrès de la Société américaine de gastro-entérologie tenu en mai 2000. La première capsule est un cylindre de 11 mm de diamètre sur 26 mm de long, d'un

Chapitre 2. La capsule endoscopique

poids de $3,451 \pm 0,35$ gramme. Ensuite, *Given Imaging* a proposé des prototypes d'une deuxième et troisième génération pour les capsules du petit intestin (*PillCam SB2*, *PillCam SB3*), dans lesquelles la source de lumière, le temps de vie de batterie, l'angle de vue et la fréquence des images (image par seconde) ont été améliorés. La société a proposé aussi deux générations d'une capsule œsophagienne à deux faces (*PillCam Eso2*), deux générations de la capsule colique (*PillCam Colón 2*) avec une fréquence d'image adaptée et une très large gamme de vue. Ensuite, le progrès technique a conduit à l'introduction de certaines versions (2ème et 3ème génération) de CE pour l'intestin grêle, l'œsophage et le côlon. En plus, plusieurs autres prototypes ont été fabriqués par des sociétés japonaises, chinoises, coréennes et américaines.

2.3. Les applications

La capsule endoscopique sans fil est devenue la première procédure du diagnostic dans certaines pathologies du système gastro-intestinal (GI) à commencer par l'œsophage jusqu'au colon. La figure 2.2 montre des exemples visuels des images extraites d'une séquence vidéo obtenue par un examen, en utilisant une capsule sans fil. Les premières capsules endoscopiques sont proposées pour l'intestin grêle, qui reste comme une boîte noire dans le système GI. Plus souvent, la capsule dans l'intestin grêle est utilisée chez les patients présentant des saignements présumés ou pour identifier les signes de la maladie de Crohn active [63]. La CE peut être utilisée également pour les patients atteints de maladie cœliaque, de syndromes de polypose et d'autres troubles de l'intestin grêle. Il y a aussi d'autres capsules spécifiques comme les capsules œsophagiennes ou coliques. Les capsules œsophagiennes peuvent être utilisées pour diagnostiquer l'œsophagite, l'œsophage de Barrett et les varices, mais on ne peut pas compter sur elles dans l'identification de la maladie gastro-duodénale [63]. Les capsules coliques peuvent être utilisées pour diagnostiquer le cancer du côlon et aussi des maladies traditionnelles du côlon. Donc, en quelques années seulement, la capsule endoscopique a évolué très rapidement, et devenait un outil inestimable pour l'examen de presque l'ensemble du tractus gastro-intestinal, comme l'intestin grêle, l'œsophage et colon, et les résultats des diagnostics ont dépassé de loin les anticipations [63].



(a)



(b)

Figure 2.2- Images du (a) estomac et (b) intestin grêle prenent par une capsule endoscopique sans fil.

2.3.1. La Capsule de l'intestin grêle

La figure 2.3(a) montre une capsule sans fil de l'intestin grêle (*Pillcam SB2*), qui est habituellement utilisée dans un cadre ambulatoire. Certains centres médicaux utilisent un régime de liquide clair pendant 24 heures avant l'étude pour améliorer la visualisation dans l'intestin grêle. Plusieurs études suggèrent que l'utilisation d'une préparation intestinale totale ou partielle



(a)



(b)



(c)

Figure 2.3-(a) capsule l'intestin grêle, (b) capsule œsophage, (c) capsule colique.

la nuit avant l'étude donne une meilleure visualisation de l'intestin grêle [64-65]. Au moment de la procédure, le système de détection est appliqué sur la paroi abdominale et relié à l'enregistreur de données porté par le patient. La capsule est activée dès qu'elle est retirée de son support magnétique. Après la prise de la capsule, les patients sont invités à tenir compte des symptômes et à surveiller les lumières de l'enregistreur de données afin de confirmer que le signal est reçu.

Les patients sont encouragés à éviter les exercices ou les activités susceptibles de provoquer le décollement des capteurs. Un régime de liquides transparents est autorisé après 2 heures et un repas léger après 4 heures. Le système d'enregistrement des données utilisables peut être déconnecté du patient après la fin de la durée de vie de la batterie. La capsule est jetable et

Chapitre 2. La capsule endoscopique

Tableau 2.1-Systèmes et spécifications de capsules sans fil.

Société	Taille (mm)	Poids (g)	Champ de vision	Image/sec	Temps vie de batterie	Résolution (pixels)
Endocapsule, Olympus American	11x26	3,5	145°	2	8 heures	512x512
PillCam SB2, Given Imaging	11x26	2,8	156°	2	8 heures	256x256
PillCam SB2EX, Given Imaging	11x26	2,8	156°	2	8 heures	256x256
Microcam, Intromedic Co Ltd	11x24	3,3	170°	3	11 heures	320x320
PillCam ESO2, Given Imaging	11x24	<4	169°	18	0.33 heures	256x256

conçue pour être excrétée. L'enregistreur de données est ensuite connecté à une station de travail pour le transfert des images acquises.

2.3.2. La Capsule œsophagienne

La capsule *PillCam ESO2*, montrée dans la figure 2.3 (b), est un système de la CE actuellement disponible pour les applications œsophagiennes. Des spécifications sur ce système peuvent être également examinées dans le tableau 2.1. Les dimensions de la capsule, la longueur d'onde de transmission, le champ de vision et la taille minimale de l'objet sont similaires au *PillCam SB2*. Cependant, la durée de vie de la batterie de la capsule est de seulement 20 minutes (contre 8-12 heures pour les capsules de l'intestin grêle), les caméras sont situées aux deux extrémités de la capsule et la capsule prend 18 images par seconde (contre 2-3 images par seconde pour la capsule d'intestin grêle). Au moment de l'examen, le patient devrait jeûner pendant 2 heures. Les patients sont équipés de trois capteurs thoraciques qui sont connectés à l'enregistreur de données. Comme dans le protocole traditionnel, le patient consomme 100 ml

Chapitre 2. La capsule endoscopique

d'eau en position debout, puis la capsule est activée en position couchée avec 20 gouttes de Siméthicone qui peut être administrée à l'aide d'une seringue. Un protocole de 5 minutes en gestion est recommandé, comprenant un enregistrement de 2 minutes avec le patient en position couchée, puis l'angle du lit est augmenté à 30 ° pendant 2 minutes, et à 60 ° pendant une minute, suivi d'une position droite jusqu'à 15 minutes pour maximiser le temps de la capsule à capturer les images, alors qu'elle traverse l'œsophage [66]. Comme le *PillCam SB2*, les images sont transmises à l'enregistreur de données et ensuite au poste de travail pour examen et interprétation via le logiciel propriétaire.

2.3.3. La Capsule colique

La deuxième version de la capsule endoscopique pour le colon (*Pillcam Colon 2*) comme montrée dans la figure 2.3 (c)) a été fabriquée pour augmenter la sensibilité pour la détection des polypes colorectaux. *Pillcam Colon 2* permet d'acquérir des images jusqu'à un maximum de 35 images par seconde et offre une couverture de 172 ° du colon. La société européenne de gastro-intestinal conseille d'utiliser *PillCam Colon 2* comme un outil fiable et sûr pour la visualisation du colon chez les patients. En bref, une méta-analyse de 8 études de la capsule colique de première génération (*PillCam Colon, Given Imaging*) a révélé une sensibilité par patient pour les polypes de toutes tailles et les résultats sont significatifs où 71 % à 68 % par rapport à la coloscopie conventionnelle [67]. En outre, les essais polyvalents de la capsule colique de la deuxième génération ont montré une sensibilité pour les polypes supérieurs ou égaux à 10 mm de 88 % à 89 %, par rapport à la coloscopie conventionnelle [68-69]. De plus amples détails sur cette technologie et ses applications potentielles peuvent être retrouvés dans le document "ASGE Report on Emerging Technology" sur la capsule endoscopique du colon [70].

2.4. Technologie

La CE est un outil du diagnostic endoscopique non invasif et ambulatoire permettant de visualiser le tractus digestif du patient et d'en localiser les éventuelles lésions. Le système de diagnostic de la capsule, montré dans la figure 2.4, comprend trois éléments :

- La capsule unique.
- Un dispositif d'enregistrement.



Figure 2.4- Système de capsule endoscopique : capsule, dispositif d'enregistrement et poste de travail.

- Une station de travail.

2.4.1. La capsule

La CE est fabriquée avec un matériau biocompatible résistant à l'action des enzymes digestives. La capsule est composée de différents éléments tels que : un capteur d'image, un système d'illumination, un système de communication (transmetteur), un système de traitement et une source d'énergie (batterie) (figure 2.5).

Capteur d'image

Deux types de capteurs d'images sont utilisés dans les capsules endoscopiques à savoir : le CCD et le CMOS. Les principales caractéristiques de ces technologies sont :

Le capteur d'image CCD (*Charge Coupled Device*) transfère une charge presque complète de pixels, créé par une conversion photoélectrique signifiant que le bruit est rare, mais un différentiel de haute tension est nécessaire pour améliorer l'efficacité du transfert augmentant

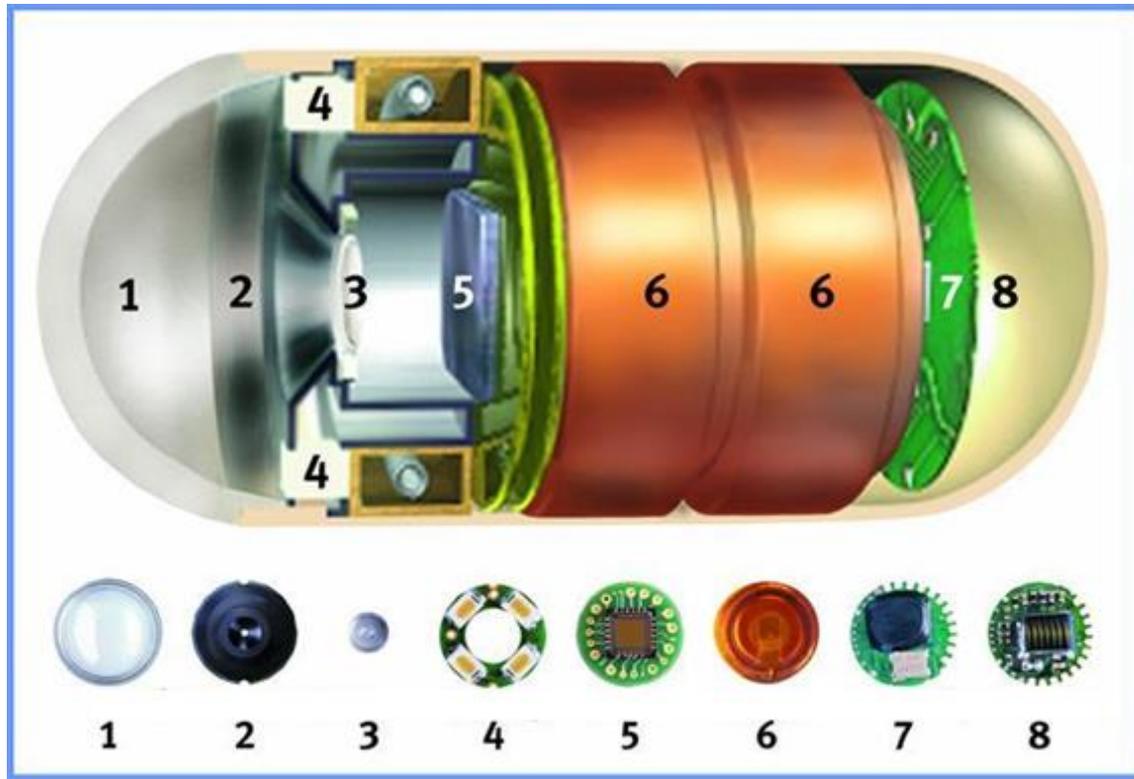


Figure 2.5- Structure de la capsule endoscopique : 1 —Dôme optique, 2 — Porte lentille, 3 — Lentille, 4 — Les LED (*Ligh Emitting Diode*), 5 — Capteur d'image CMOS, 6 — Batteries, 7 — Transmetteur ASIC, 8 — Antenne.

aussi la consommation d'énergie, et c'est pour cette raison que, actuellement, la plupart des capsules endoscopiques n'utilisent pas ce capteur.

Le capteur d'image CMOS (*CoMplementary Oxide Semiconductor*) convertit la charge en une tension pour chaque pixel. Il utilise une matrice de pixels pour convertir la lumière en des signaux électriques. Ce signal est faible et doit être amplifié à un niveau utilisable, où chaque pixel possède son propre circuit amplificateur. Le capteur CMOS contient un nombre de connecteurs plus faibles, une fiabilité accrue, une consommation d'énergie réduite et il est plus compacte.

Dans les capsules endoscopiques, en avant du capteur d'image (CMOS ou CCD), on trouve une lentille asphérique qui est focalisée manuellement sur le dôme optique, de sorte que l'intestin se trouve contre le dôme optique. Le choix des lentilles permet un examen plus efficace de l'intestin. Les améliorations de la qualité d'image peuvent être obtenues grâce à une attention particulière au positionnement de la lentille par rapport au capteur. La position de la pupille

Chapitre 2. La capsule endoscopique

d'entrée de la lentille optique doit coïncider avec le centre de la courbure de la fenêtre du dôme optique pour empêcher toute lumière réfléchie d'entrer dans le capteur d'image [71, 72].

Système d'illumination

Le défi dans le développement d'un système d'éclairage LED pour la CE est l'obtention d'un éclairage uniforme [73] sur l'objet. Ceci est difficile pour plusieurs raisons :

- Le trajet optique d'une source de lumière LED n'est pas parallèle à l'axe optique des lentilles d'imagerie adjacentes ;
- Le motif lumineux de la LED dépend sensiblement du courant, de l'emplacement et des angles de projection ;
- L'intestin observé n'est pas plat ;
- L'intensité lumineuse d'une LED est différente selon les différents angles.

En général, la répartition de la lumière spatiale n'est pas uniforme sur le plan de l'objet devant la capsule, ce qui entraîne une réduction de la qualité d'image. Dans certaines régions du plan de l'objet, la luminance est très petite et, dans d'autres régions, elle pourrait être très grande. Afin de résoudre ce problème, les positions et les angles de plusieurs LED ont été agencés pour améliorer l'uniformité lumineuse de la capsule. En raison des contraintes d'espace dans la capsule, les LED doivent être petites et doivent fournir une intensité lumineuse suffisante pour obtenir des images de bonnes qualités. Leurs intensités lumineuses devraient également pouvoir fournir suffisamment de lumière pour le pire scénario où la lumière nécessaire pour illuminer le côlon est très grande : dans le cæcum et dans le côlon ascendant. La puissance requise par les LED dépend du nombre et du type de LED utilisé.

Transmetteur et récepteur

La CE a besoin d'un système de transmission efficace qui ne consomme pas beaucoup d'énergie, et qui peut transmettre un signal qui pourra traverser le corps humain. La transmission de ce signal est aussi un défi pour les chercheurs [74]. La transmission radio fréquence (RF) doit fonctionner dans une variété d'environnements et de positions qui peuvent changer avec le temps, car le tissu du corps humain atténuera le signal de manière variable lorsque la capsule traversera le système GI.

Chapitre 2. La capsule endoscopique

Il existe plusieurs exigences critiques pour l'émetteur dans la CE : la taille, la consommation d'énergie et le débit de données. La petite taille est essentielle, car il y a très peu de place pour l'intégrer dans la capsule avec les autres composants. Le matériel de transmission doit être robuste afin de résister aux coups et aux bosses du mouvement normal du corps humain. La connexion entre l'émetteur et l'antenne doit être très courte pour minimiser les pertes et pour maintenir l'impédance constante. La longueur de l'antenne doit être suffisamment petite pour s'adapter à la capsule. Cependant, si l'antenne compacte dans la capsule est trop petite, sa bande passante à basse fréquence ne sera pas assez large et son efficacité de rayonnement ne peut pas être très élevée. En outre, afin de transmettre les données d'image à haute résolution en temps réel et en raison de l'effet du tissu du corps humain sur les caractéristiques de propagation du signal, la fréquence de transmission peut légèrement se décaler en cours de fonctionnement [75]. Donc, une large bande passante et une fréquence de transmission élevée doit être utilisée avec un récepteur accordable. Malheureusement, une antenne avec une fréquence de transmission plus élevée peut provoquer une absorption de rayonnement plus élevée, en particulier dans la partie la plus reculée de l'intestin qui dégrade les performances de la liaison de communication dans le corps humain. En plus de l'énergie absorbée par les tissus humains, de nombreux autres problèmes existent tels que la consommation d'énergie, les changements de positions et les orientations de la source de rayonnement, qui influencent les caractéristiques de la communication [76]. Par conséquent, une sélection soignée de l'émetteur, de l'antenne et de la fréquence pour une utilisation *in vivo* est très importante. Parmi les fréquences ISM (industrielles, scientifiques et médicales), la fréquence 1,9 GHz est fréquemment utilisée pour la transmission, et les plus petits émetteurs disponibles dans le commerce fonctionnent dans cette gamme.

L'alimentation et l'encapsulation

Les CE commercialisées utilisent de petites batteries d'oxyde d'argent (généralement, il y en a deux), qui représentent plus de la moitié du poids de la capsule. Dans toutes les applications de la capsule endoscopique, il est impératif de tenir compte des compromis de la durée de vie de la batterie et de la toxicité de la technologie de la batterie. Un choix judicieux de composants et une conception électrique efficace sont importants pour améliorer la performance des capsules. Il existe seulement certains types de piles, disponibles dans le commerce, qui peuvent s'insérer

Chapitre 2. La capsule endoscopique

dans l'espace limité de la capsule. Donc, l'alimentation électrique doit être conçue de manière très efficace. La batterie CR 1/3N fournit un courant continu suffisant pour permettre le fonctionnement fiable de tous les éléments électroniques à l'intérieur de la capsule. En plus de la tension de fonctionnement, le courant de décharge continu, le poids faible et les dimensions petites sont aussi des facteurs décisifs lors du choix d'une alimentation électrique efficace. Donc, un paramètre important à considérer est la résistance interne de la batterie, car elle doit fournir suffisamment de courant de décharge continue de 28-30mA pour les composants électroniques et une impulsion de 400 ms/80mA pour déclencher le microchauffage simultanément. Plus la résistance interne est basse, moins la batterie est résistante à la livraison des pointes d'alimentation nécessaires. En outre, plus la batterie est petite (c'est-à-dire son diamètre), plus la résistance interne est élevée.

Pour le premier prototype de CE, son enveloppe extérieure est une couche de gélatine dure de TorpacInc (Fairfield, NJ, USA) qui se dissout très rapidement (2-3 minutes) dans l'eau. Cette capsule est dimensionnée pour les tests sur les animaux. Dans le prototype actuel, l'accent est mis sur la démonstration de la stabilisation et de la séparation, et le test suppose que l'enveloppe extérieure permette d'atteindre l'organe ciblé.

2.4.2. Le dispositif d'enregistrement

Le dispositif est complété par un ensemble de capteurs placés sur la peau du malade au niveau de l'abdomen, qui recueille le signal émis par la capsule, la figure 2.6 montre le capteur et l'outil d'enregistrement.

Data Recorder : le signal est transmis à un enregistreur de haute fréquence et un disque dur, contenu dans un boîtier. Les images sont ensuite transférées à une station de travail.

SensorArray : est un ensemble d'antennes réceptrices (8 capteurs plats de 40 mm de diamètre) réparties sur l'abdomen du patient selon des repères anatomiques simples. Les électrodes sont maintenues à la peau par des manchons adhésifs à usage unique.

Le bloc-batterie : fournit l'énergie au Data Recorder TM et comporte 5 accumulateurs au nickel métal de 1,2 V rechargeable.

Le Recorder Belt : est une ceinture adaptable munie de bretelles permettant de loger le Data Recorder et le bloc batterie.

2.4.3. La station de travail

La station de travail est un ordinateur dédié au téléchargement et à l'analyse des images enregistrées, figure 2.7. Elle se compose de :

- Une unité centrale ;
- Un écran plat ;
- Une interface de connexion avec l'enregistreur de données
- Le Data Recorder TM ;
- Un graveur interne de CD-ROM ;
- Un disque dur externe de stockage de 160 Go ;
- Une imprimante couleur ;

Depuis 2005, il est possible de voir les images de l'intestin grêle en temps réel grâce au système « Rapid Access » (Reporting And Processing of Images and Data) de Given Imaging, connecté directement au boîtier d'enregistrement porté par le patient.



Figure 2.6- Capteurs et boîte d'enregistrement

2.5. Les défis scientifiques

La capsule endoscopique sans fil est appliquée de plus en plus non seulement en raison de sa commodité et qu'elle soit sans douleur, mais aussi en raison de ses effets cliniques. En comparant la CE à l'endoscopie traditionnelle (par la fibre optique), les apparences du tractus GI observés avec la CE étaient généralement similaires à celles de l'endoscopie traditionnelle, et la capsule était nettement supérieure à l'endoscopie traditionnelle dans l'identification des sources de saignement dans l'intestin grêle [77]. La capsule est également significativement supérieure à l'endoscopie traditionnelle dans le diagnostic des anomalies du tractus GI ainsi que dans les saignements [78-79]. Dans l'ensemble, la CE est un moyen efficace pour faire des investigations dans le tractus GI, avec un rendement élevé et avec une grande fiabilité. En outre, c'est la seule méthode d'imagerie qui peut fournir des images de couleur de l'intestin grêle inférieur. De plus, sans considérer le coût, les patients préfèrent la CE à l'endoscopie traditionnelle.

Néanmoins, il existe des limites à la CE. Tout d'abord, la qualité de l'image n'est pas aussi bonne que celle de l'endoscopie par la fibre optique. La fréquence d'images avec la capsule est inférieure (2-18 vs 25 images /sec). La résolution de l'image de 256×256 pixels n'est pas satisfaisante et l'intensité de la lumière ne peut pas encore répondre aux exigences nécessaires pour un changement complet vers la CE. La plupart des capsules endoscopiques qui existent dans le marché sont alimentées par des batteries. Ces capsules ne peuvent travailler que pendant 8 heures, en raison de l'alimentation électrique limitée. Néanmoins, pour un patient atteint d'une maladie du tractus GI, il faut beaucoup plus que 8 h pour le diagnostiquer ; sinon, l'intestin grêle reste toujours non examiné lorsque la batterie est épuisée. Enfin, le déplacement de la CE est passif et sa localisation ne peut pas être contrôlée, de sorte que les lésions ne peuvent être examinées à plusieurs reprises. Par conséquent, la CE est encore une technologie immature, et les améliorations techniques doivent être encouragées.

Néanmoins, la technologie de la capsule endoscopique sans fil a été appliquée en clinique avec succès, en particulier dans le diagnostic de l'intestin grêle.

Le temps de travail, la résolution de l'image et la fréquence d'images restent absolument limités par l'alimentation électrique. On peut aussi améliorer la CE en améliorant les composants qui consomment beaucoup d'énergie tel que le capteur d'image, l'unité de traitement d'images et le module RF. Maintenant, la société *Given Imaging* et la société *ChineseHitron* n'utilisent plus les

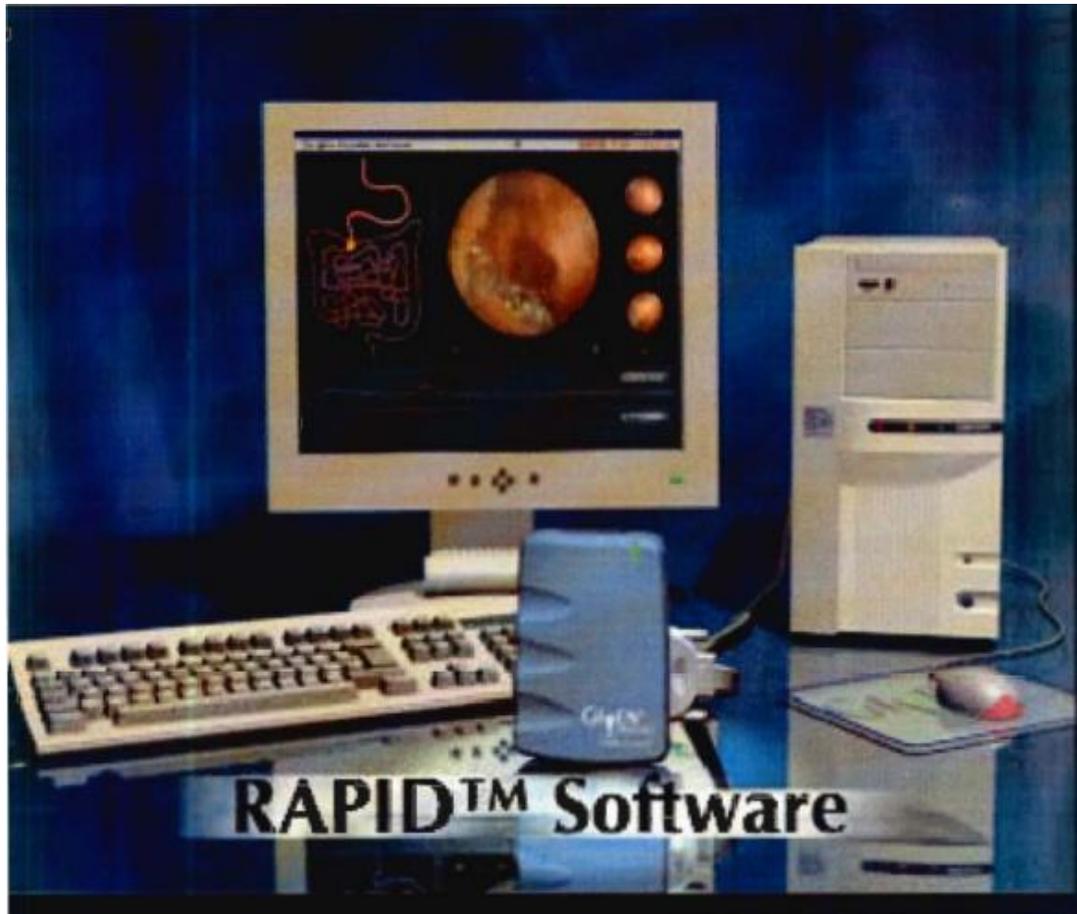


Figure 2.7- Poste de travail

circuits à usage général pour concevoir la CE. Ils ont personnalisé leurs puces exclusives qui intègrent le module de traitement d'image et le module RF. Ainsi, la majeure partie du circuit diminue de manière significative et la consommation d'énergie, dans la CE, diminue considérablement. Par conséquent, le temps de travail de la CE est étendu à plus de 15 heures. Cependant, ce n'est pas suffisant. Si on améliore la résolution des images et la fréquence des images, la consommation de puissance augmente de manière significative et le temps de travail diminue de façon spectaculaire. Par conséquent, la limitation de l'alimentation est toujours un point noir dans le système de la CE [73].

Une autre tendance de cette nouvelle technologie est la fréquence d'images élevée et les images à haute résolution. Les organes GI vivants se déplacent en permanence, de sorte qu'une fréquence d'images de plus de 2 images/sec est évidemment nécessaire pour diagnostiquer les détails des organes du tractus GI. Maintenant, certains systèmes de la CE ont été développés qui

Chapitre 2. La capsule endoscopique

peuvent transmettre des images avec une fréquence de 30 images/sec en format vidéo NTSC [80]. D'autre part, la résolution d'image de 256×256 pixels n'est pas satisfaisante par rapport à l'endoscopie câblée, et la résolution de l'image ne peut plus être plus élevée. Cependant, la fréquence d'image élevée et la haute résolution d'image nécessitent en conséquence beaucoup plus d'énergie qui est au-delà de ce que toute pile peut fournir. Afin de réduire les données d'image pour la haute résolution et la fréquence élevée de l'image, de nouveaux algorithmes de compression ont été développés pour la CE [81, 82]. Cependant, l'algorithme de compression complexe ne convient pas aux capsules endoscopiques de faible puissance et ces effets sont limités. La transmission de puissance sans fil peut être une bonne solution pour ce problème, mais ce n'est pas facile, et c'est donc un grand défi technique.

Une autre tendance est la CE active, à savoir, la capsule robotique, un prototype de ce type de capsule est montré sur la figure 2.8. La CE active peut facilement intégrer de nouvelles fonctions telles que la biopsie et la délivrance de médicaments. Maintenant, dans le système existant, lorsque la CE est avalée, elle se dirige vers l'avant avec le péristaltisme naturel du tractus GI, de sorte que la locomotion de CE est passive. En conséquence, la position ne peut pas être contrôlée, ce qui constitue le principal inconvénient de la technologie de la CE. En raison de la

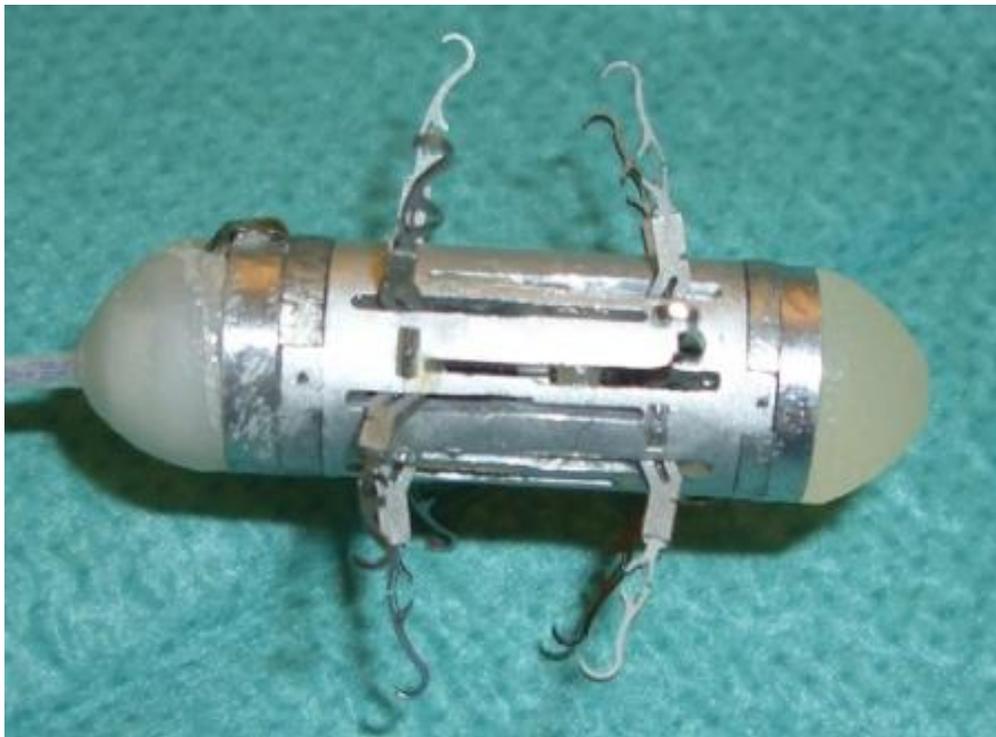


Figure 2.8- Un prototype d'une capsule robotique.

locomotion passive de la CE, on ne peut pas envisager des prises de vue d'une maladie avec des orientations différentes et à plusieurs reprises. En outre, la locomotion passive est un point noir de la CE avec diverses fonctions, y compris la biopsie et la délivrance de médicaments. La CE active sera donc la prochaine génération de la CE passive. Néanmoins, la locomotion de la CE active dans le tractus GI est un autre grand défi, en plus de celui de l'alimentation électrique. Le tractus GI est un environnement très difficile pour la locomotion active. Le tissu est normalement mou et viscoélastique, et il est très difficile pour la CE active en mouvement de ne pas nuire au tissu. La morphologie du GI varie selon le tractus GI et, par conséquent, la direction du mouvement sera de haut en bas, en avant et en arrière. Certains chercheurs ont intégré un actionneur magnétique dans la CE et ont utilisé un champ magnétique solide vers l'extérieur pour contrôler le déplacement de la CE [83-85]. Ce type de CE autopropulsé nécessite un système d'actionnement complexe et est difficile à utiliser.

2.6. Conclusion

La capsule endoscopique (CE) a ouvert une nouvelle ère dans le diagnostic endoscopique des gastro-entérologies et a marqué une étape importante dans l'évolution de l'examen endoscopique du tractus gastro-intestinal sans besoin de sédation, et permet d'éviter les risques impliqués par l'endoscopie conventionnelle. La CE a démontré son rendement élevé dans le diagnostic des pathologies multiples du tractus gastro-intestinal tel que les saignements gastro-intestinaux obscurs, la maladie cœliaque et les petites et grandes tumeurs intestinales.

Dans ce chapitre, on a fait une description générale sur les capsules, leur composition, leur performance et aussi sur ces limites : comme la locomotion passive, incapacité à effectuer une biopsie et les limites énergétiques qui restent un défi à surmonter par les chercheurs en éliminant, ou au moins en réduisant, les inconvénients de la CE.

En conclusion, puisque les générations futures de la CE sont destinées à transmettre une résolution accrue, une fréquence d'images élevées et à avoir une durée de vie de batterie plus longue, elles nécessitent donc une compression vidéo efficace à un coût de calcul aussi faible que possible. Le DVC est un candidat fort pour répondre aux exigences techniques imposées par la capsule endoscopique sans fil, offrant un encodage à faible coût, une évolutivité et une efficacité de compression élevée [86].

Chapitre 3

Codage vidéo distribué basé sur une base de données

3.1. Introduction

Dans ce chapitre, on va discuter les deux techniques de codage DVC, utilisant une base de données (BDD) pour générer l'information adjacente (SI), que nous avons proposées. Les deux contributions proposées utilisent un codeur très simple avec seulement une trame WZ, contrairement à d'autres systèmes DVC qui utilisent des trames WZ et des trames clé. La première contribution (DVC-HDC) [87] utilise les coefficients DC comme une information de hachage et la deuxième contribution (DVC-HSE) [88] utilise une image sous-échantillonnée comme une information de hachage. On a proposé aussi un codage pour les composantes couleurs (chrominance) de l'image de la CE. Ce chapitre explique le déroulement du codage et du décodage des trames vidéo de la capsule endoscopique et discute les résultats obtenus dans le cas des deux contributions.

3.2. Similarité entre images de capsules endoscopiques

Avant de proposer une architecture de codage pour les images de la CE, il faut d'abord examiner les relations entre les images endoscopiques. Les organes dans les corps humains (l'œsophage, l'estomac ou l'intestin) sont très similaires. Cette corrélation, entre les images, permet de proposer un système de compression pour les vidéos de la CE. Le codec conventionnel d'une vidéo exploite la corrélation temporelle codée par un codeur inter, et la corrélation spatiale d'une image qui est codée par un codeur intra. Dans le contexte de DSC, nous pouvons modifier l'utilisation de la corrélation pour inclure d'autres sources de corrélation entre les images, qui peuvent être incluses dans la définition de " l'information adjacente disponible au décodeur ".

On considère que nous avons deux capsules qui transmettent simultanément à partir de deux corps. X est le signal du premier corps et Y est celui du second. X et Y sont des signaux très corrélés si la synchronisation est assurée (image de la même région). En comparaison avec le codage vidéo WZ proposé dans la littérature, X et Y ne sont pas deux trames voisines de la même séquence mais deux sources différentes, et en plus très corrélées.

Pour réaliser un système WZ pratique, nous proposons un autre changement qui est : Y ne provient plus d'un corps réel, mais d'une autre source comme une base des données (BDD). Le défi de ce système serait de maintenir la meilleure synchronisation possible entre les deux séquences vidéo et de trouver la meilleure information adjacente possible. La figure 3.1 résume notre idée. Considérons le scénario suivant :

- 1- Dans (a), nous avons deux humains qui avalent deux capsules en même temps. Si l'on considère que le mouvement de la capsule est synchronisé, les deux capsules seront deux sources corrélées pouvant être des candidates idéales pour un codage distribué symétrique. La corrélation est due au fait que l'anatomie humaine est la même.
- 2- Dans (b), nous avons un scénario asymétrique où un humain est utilisé comme une information adjacente pour décoder l'autre.
- 3- Dans (c), nous remplaçons le second humain par une base de données (BDD) qui propose une information adjacente.

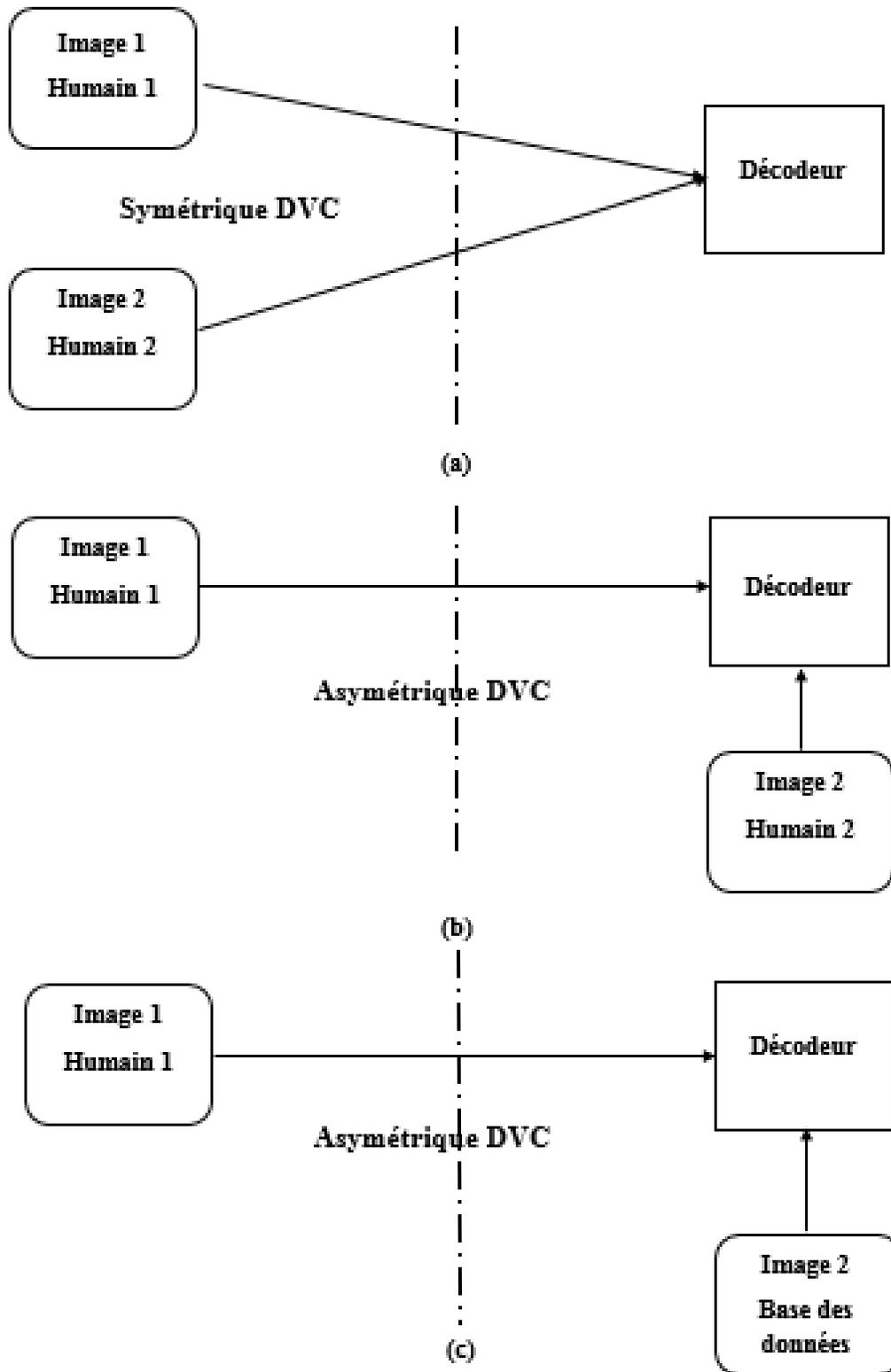


Figure 3.1- Exemple de notre idée

3.3. Utilisation des coefficients DC comme une information de hachage

Parmi les méthodes de compression proposées dans la CE, il y a le codage sans perte basée sur JPEG-LS [89-90]. Il donne une très bonne qualité d'image, mais la compression n'est pas élevée. Dans [91], des algorithmes de compression d'image basés sur la DCT sont proposés pour la CE. Ces algorithmes [91-93] sont complexes et donnent une mauvaise qualité d'image. Pour proposer une nouvelle architecture pour la CE, quelques critères doivent être respectés :

- 1- le fonctionnement de la capsule devrait avoir une faible consommation d'énergie, car la durée de vie de la batterie est limitée. Par contre, les algorithmes de codage vidéo sont très complexes (tels que l'estimation et l'interpolation du mouvement) et ne peuvent être effectués dans la capsule. Nous nous concentrons donc sur des algorithmes d'une faible complexité. Cela donnera la possibilité d'ajouter des fonctionnalités supplémentaires (par exemple, la robotique, le contrôle de vitesse [94-96], etc.) dans les capsules disponibles.
- 2- La taille d'une capsule typique est limitée à 28×11 mm [97]. Ainsi, l'espace dans la capsule est un problème critique qui limite l'utilisation de la mémoire. La mémoire consomme plus de surface du silicium et plus de puissance. Ici, nous nous concentrons sur des algorithmes qui ne nécessitent pas beaucoup de mémoire.
- 3- Enfin, l'algorithme de compression doit également réduire les données afin que l'émetteur puisse envoyer des images à une fréquence d'environ 2 images par seconde en utilisant sa bande passante limitée.

Le codage, même en DVC, dans la CE n'est pas comme dans d'autres applications, car nous ne pouvons pas utiliser l'estimation de mouvement. Ceci est dû au fait que le mouvement est très complexe et il n'y a aucune stabilité de la capsule à l'intérieur du corps. Dans [98], les auteurs proposent un codec DVC pour la CE et ils ont utilisés des trames clé et des trames WZ. La trame clé est codée par un codec Motion JPEG (MJPEG) et la trame WZ est codée en deux étapes. La trame WZ est sous-échantillonnée et codée par MJPEG et transmise comme hash. Lorsque cette trame est décodée, le décodeur l'utilise avec la trame clé pour créer la SI et décodeur la trame WZ complète ; qui est codée par un codeur SW. Dans notre DVC (DC-HDC) [87], nous utilisons uniquement des trames WZ (pas de trames clé), mais nous encodons différemment les coefficients DC et AC. Le schéma du système DVC-HDC proposé est donné à la figure 3.2.

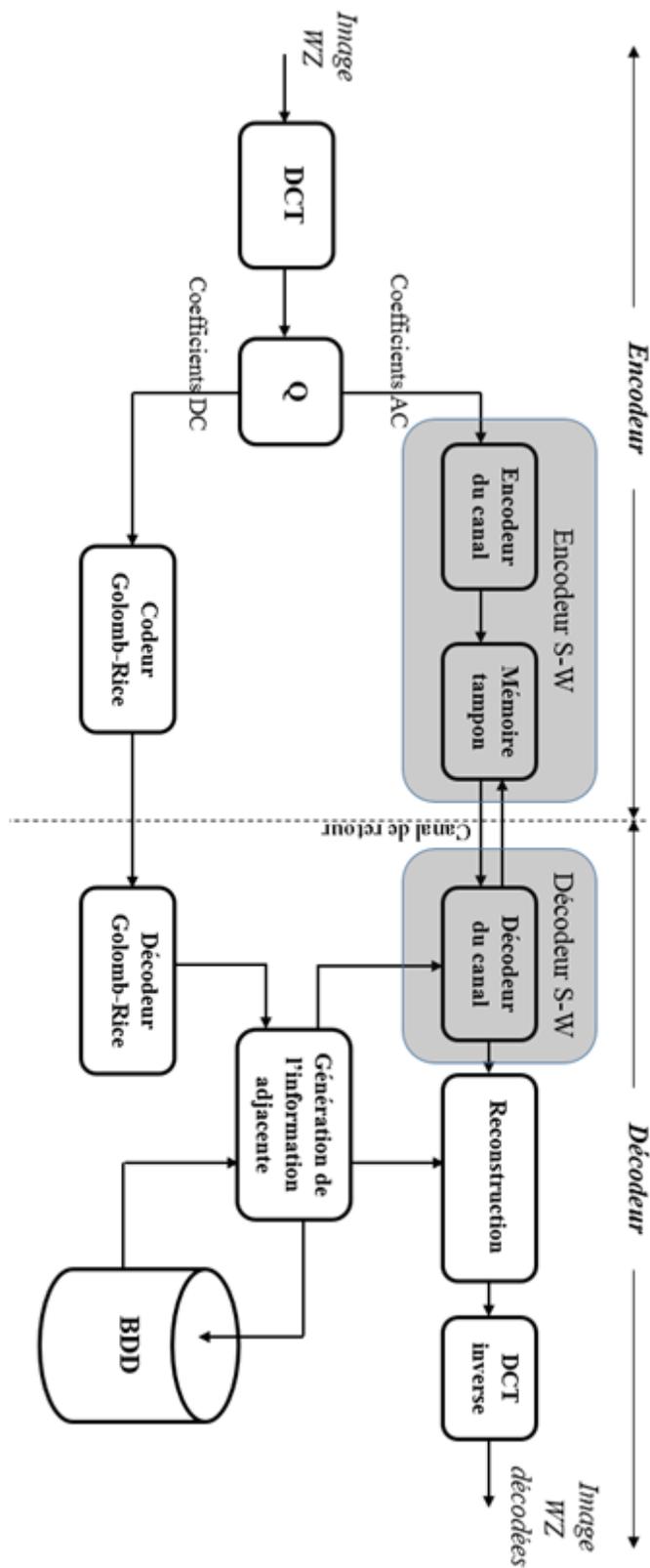


Figure 3.2-L'architecture de DVC proposé pour la capsule endoscopique.

3.3.1. La construction de la base de données

L'organisation d'une BDD passe par plusieurs étapes, au début, nous rassemblons des séquences vidéo, provenant de CE, du système gastro-intestinal(GI) sain ou malade. Puis, nous classifions chaque image par sa localisation dans le système GI comme l'intestin grêle, l'estomac ou le colon. Ensuite, la construction d'une BDD se fait par l'utilisation d'un algorithme de sélection qui fonctionne comme suit :

- Nous plaçons une première image dans la BDD
- Ensuite, chaque nouvelle image est comparée avec les autres images qui sont déjà dans la BDD : si l'image est très différente des images de la BDD, nous l'ajoutons dans la BDD, sinon elle est rejetée.

On utilise le PSNR (*Peak Signal to Noise Ratio*) pour décider si l'image ressemble ou non aux autres images de la BDD. Pour mettre l'image dans la BDD, il faut la comparer avec toutes les images de la BDD, si le PSNR est inférieur à un PSNR seuil de 24dB, alors on met l'image dans la BDD sinon nous la rejette. Nous avons choisi la valeur du PSNR seuil après plusieurs tests pour obtenir un compromis entre une BDD qui contient suffisamment d'images pour produire une bonne qualité de SI et une taille petite de BDD, pour réduire la complexité. Donc, la taille de la BDD est contrôlée par le PSNR seuil, si le seuil est grand la taille de la BDD va être petite et inversement si le seuil est petit la taille de la BDD devient grande.

$$PSNR = 20 \log_{10} \frac{255}{\sqrt{\frac{1}{M \times N} \sum_{n=1}^N \sum_{m=1}^M (x_{m,n} - \hat{x}_{m,n})^2}} \quad (3.1)$$

Où M , N sont les dimensions d'image et $x_{m,n}$, $\hat{x}_{m,n}$ sont l'image et l'image de la BDD.

Il existe deux méthodes de recherche dans la BDD : par l'utilisation de la recherche exhaustive ou par la recherche orientée (comme la recherche par arbre). La recherche exhaustive permet d'obtenir la meilleure image dans la BDD qui est la plus similaire à l'image source. Cependant, la taille de la BDD rend le décodeur très complexe, car le nombre des opérations pour sélectionner une image à partir de la BDD est très grand. Donc, pour effectuer la recherche et la sélection d'une image SI dans la BDD, on utilise les coefficients DC comme information de hachage. Dans le système DVC-HDC [87], on a préparé les coefficients DC des images de la

Chapitre 3 : codage vidéo distribué basé sur une base de données

BDD et on les a mis dans un tableau. Lorsque les valeurs de l'information de hachage arrivent au décodeur elles sont comparées directement avec les coefficients DC des images de la BDD. Ceci nous permet d'éliminer le besoin de calculer la DCT pour chaque bloc de l'image de la BDD et permet donc d'augmenter la vitesse de la recherche.

On peut aussi réduire la complexité du décodeur si on utilise une recherche qui fonctionne avec le système de localisation dans la capsule endoscopique. Dans ce cas, il faut organiser les images de la BDD en fonction de leur similitude et aussi il faut partitionner la BDD en sous-groupes qui contiennent : des images d'estomac, des images de l'intestin grêle ou des images du colon. Ensuite, grâce au système de localisation de la CE, on peut sélectionner un seul sous-groupe pour rechercher la SI. Cette recherche permet de réduire le temps de recherche dans la BDD et réduire aussi la complexité du décodeur.

3.3.2. L'encodeur

Dans notre système, nous convertissons les images qui arrivent à partir du capteur CMOS de l'espace de couleur RVB à YUV en utilisant l'équation (3.2). Dans l'espace YUV, la composante Y représente l'intensité (luminance), alors que les composantes U et V contiennent des informations sur les couleurs (chrominance). À partir des tests effectués dans [96], les valeurs des composantes U et V ne varient pas beaucoup avec les positions et restent presque constantes dans le même plan, contrairement à l'espace RVB où elles varient avec les positions. Il est également à noter que cette transformation ne dégrade pas théoriquement la qualité de l'image ; aucune information n'est perdue. Le petit contenu d'information dans U et V permet d'utiliser moins de bits pour les coder, sans perdre aucune information.

$$\begin{aligned} Y &= \left(\frac{(66 \times R + 129 \times G + 25 \times B + 128)}{256} \right) + 16 \\ U &= \left(\frac{(-38 \times R - 74 \times G + 122 \times B + 128)}{256} \right) + 128 \\ V &= \left(\frac{(112 \times R - 94 \times G - 18 \times B + 128)}{256} \right) + 128 \end{aligned} \quad (3.2)$$

En outre, la plupart des capteurs d'image disponibles dans le commerce sont maintenant équipés d'un convertisseur RGB-YUV. Le schéma proposé vise à tirer parti du convertisseur sur puce, c'est-à-dire qu'il n'est pas nécessaire d'implémenter les équations de la transformation à

Chapitre 3 : codage vidéo distribué basé sur une base de données

l'intérieur du compresseur d'image, ce qui permet d'économiser de l'espace et la consommation électrique du compresseur.

3.3.2.1. Transformation cosinus discret

Dans notre système [84], les bandes DC et AC générées par la DCT sont codées séparément : les coefficients DC sont codés par le codeur Golomb-Rice et les coefficients AC sont codés par le codeur SW. Soit X l'image provenant d'un capteur CMOS. La transformée DCT entière de H.264 [34] est appliquée sur un bloc de 4x4 pixels de cette image, selon l'équation suivante :

$$X_{DCT} = CXCT^T \quad (3.3)$$

Où C est une matrice définie comme suit :

$$C = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix} \quad (3.4)$$

La transformée DCT proposée dans H.264 peut être implémentée simplement en utilisant les opérations de décalage et de l'addition, parce qu'elle contient seulement des valeurs entières. Pour les valeurs de X_{DCT} , on applique une quantification en utilisant l'équation suivante :

$$X_{DCT}^Q = \text{sign}(X_{DCT}(i,j)) \cdot [(|X_{DCT}(i,j)|M(Q_M,r) + f \cdot 2^{(15+Q_E)}) \gg (15 + Q_E)]. \quad (3.5)$$

Q_M et Q_E sont reliés au paramètre de quantification Q_P . Dans notre cas $Q_P = 1$:

$$Q_M = Q_P \text{ mod } 6 \quad (3.6)$$

$$Q_E = Q_P / 6$$

La matrice M est définie comme :

$$M = \begin{bmatrix} 13107 & 5243 & 8066 \\ 11916 & 4660 & 7470 \\ 10082 & 4194 & 6554 \\ 9362 & 3647 & 5825 \\ 8192 & 3355 & 5243 \\ 7282 & 2893 & 4559 \end{bmatrix} \quad (3.7)$$

La valeur du paramètre r est liée à la position des coefficients transformés du bloc X_{DCT} :

- $r=0$ pour $(i, j) = \{(0,0), (0,2), (2,0), (2,2)\}$,
- $r=1$ pour $(i, j) = \{(1,1), (1,3), (3,1), (3,3)\}$,
- $r=2$ pour ailleurs.

Le paramètre f définit la zone morte de la quantification et prend les valeurs entre 0 et 0.5. Cependant, dans notre cas, nous avons pris $f = 1/3$, car cela améliore les résultats. Tous les coefficients des blocs d'images sont classés sur un tableau où la première couche représente la valeur DC et les autres coefficients sont les valeurs AC.

Les coefficients issus de cette transformation, X_{DCT}^Q , sont regroupés dans le tableau des 16 bandes de fréquences ($k = 1, \dots, 16$). Les coefficients DCT de l'image sont codés en deux étapes : dans la première étape, les coefficients DC sont codés par le codage prédictif et Golomb-Rice et dans la deuxième étape, les coefficients AC sont codés par le codeur SW.

3.3.2.2. La quantification

Le pas de quantification dans notre système varie d'une image à l'autre, et pour obtenir une meilleure qualité d'image à la reconstruction, le décodeur a besoin de connaître les valeurs maximales absolues de chaque bande de fréquences. Pour cette raison, l'encodeur envoie ces valeurs au décodeur, cette transmission correspond au maximum à 240 bits par image, si on considère que chaque bande de fréquences a besoin de 16 bits. Ensuite, les coefficients subissent une quantification scalaire où le pas de quantification dépend de la valeur maximale V_k des coefficients de la bande k . Alors le pas de quantification de chaque bande codé par M_k niveaux est calculé par :

$$\begin{aligned}\Delta_{DC} &= |V_k| / (M_k - 1) \\ \Delta_{AC} &= 2|V_k| / (M_k - 1)\end{aligned}\tag{3.8}$$

Le pas de quantification change en accord avec la matrice de quantification. Les matrices 4 x 4 de quantification dans la figure 1.12 proposée par [37] contiennent des valeurs de $M_k = 2^b$ où b est le nombre de plan de bits utilisés pour la bande k et la valeur 0 signifie que cette bande de fréquences n'est pas transmise vers le décodeur.

3.3.2.3. Codage Golomb-Rice

Pour la partie DC, un codage prédictif est appliqué sur les coefficients DC et le résultat est mis dans un registre. Ensuite, on utilise un codage de Golomb-Rice. Ce codage génère des codes optimaux avec des longueurs variables. Le codage de Golomb-Rice, proposé par [99], est simple et il est facile à implémenter dans le matériel et donne un taux de compression très élevé proche de celui de Golomb [100]. On utilise ce type de codage pour coder les erreurs de prédiction dX des coefficients DC (équation 3.9).

$$dX = X(n+1) - X(n) \quad (3.9)$$

Le codeur Golomb-Rice, en général, code les valeurs entières non signées, mais dans notre cas, les valeurs prédictives sont positives et négatives. On rend toutes les valeurs positives par l'équation (3.10). Les valeurs m_{dX} codées par le codage Golomb-Rice comme dans [99]. Donc, le codeur Golomb-Rice génère pour chaque valeur un code, composé d'un code unaire et d'un code binaire. Ce code est transmis ensuite vers le décodeur.

$$\begin{cases} m_{dX} = 2 * dX & si(dX > 0) \\ m_{dX} = 2 * dX - 1 & si(dX < 0) \end{cases} \quad (3.10)$$

Théoriquement, la valeur la plus grande des coefficients DC est 2048, avec la quantification (lorsqu'on utilise la matrice de quantification $Q8$, on divise par 16) la valeur devient 128 et avec les valeurs négatives qui se trouvent dans un intervalle $[-128 \ 127]$, alors les valeurs m_{dX} sont toutes positives dans l'intervalle 0 à 255. Le codage se déroule comme dans l'article [99]. La longueur maximale des codes choisis est $g_{limit} = 26$.

Tous les coefficients DC de l'image, codé par le codage prédictif et Golomb-Rice, sont transmis vers le décodeur pour les utiliser dans la génération de SI.

3.3.2.4. Codage Slepian-Wolf

Dans la deuxième partie, l'encodeur code les coefficients AC par le codeur SW comme dans la figure 3.2. Les symboles AC quantifiés sont décomposés en plans de bits et passés à un codeur de canal (LDPC Accumulateur (LDPCA)). Comme dans [99], notre système ne code pas les coins noirs de l'image de la CE, donc on n'a pas besoin de coder 4096 valeurs par bande, mais

Chapitre 3 : codage vidéo distribué basé sur une base de données

seulement 3762 valeurs. Alors, on utilise une matrice H d'une taille 3762 et avec un bloc de bits syndrome d'une taille 57 bits, pour chaque incrémentation de taille du codeur LDPCA.

La plupart des systèmes DVC actuels incluent un canal de retour utilisé par le décodeur. L'objectif principal de ce canal est d'informer l'encodeur pour transmettre les bits de syndrome vers le décodeur, lorsque le message n'a pas pu être reçu convenablement. Ce processus permet un contrôle direct du débit binaire, mais le canal du retour rend le processus de décodage plus complexe et plus lent [101,102]. Un moyen simple fut trouvé pour pallier à ce problème. Il s'agit simplement de forcer l'encodeur de WZ à envoyer continuellement des bits de syndrome (pour chaque période de temps T) jusqu'à ce qu'il reçoive un message (d'un ou plusieurs bits) du décodeur indiquant que le décodage est correctement effectué avec les derniers bits de syndrome reçu. Ceci permet également à l'encodeur de commencer la transmission d'un autre message.

La conversion des coefficients DCT quantifié en plan de bits est une opération nécessaire pour le codeur LDPCA binaire. Les valeurs des coefficients DCT de l'image sont converties en bits, qui seront réorganisés en plans de bits (Figures -3.3) à partir des bits MSB suivi par les bits de poids inférieur jusqu'aux bits de poids le plus faible (LSB).

Les plans de bits seront de taille 3762 bits chacun prêt à être codé par le codeur LDPCA. Chaque

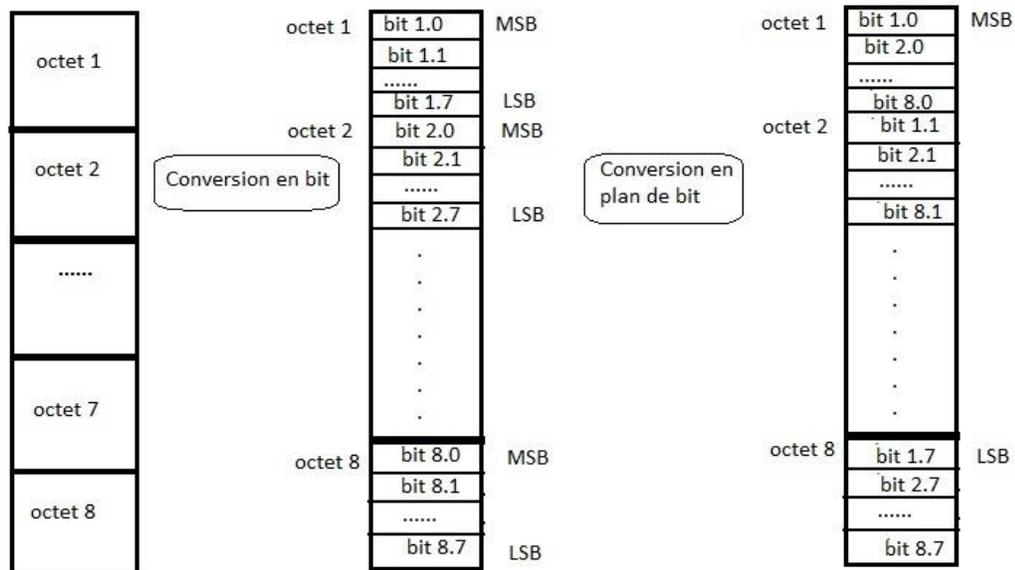


Figure 3.3- Conversion de l'image en plans de bits

paquet X du plan de bits est ensuite multiplié par la matrice de parité H à poids faible, de taille 3762 pour obtenir les syndromes S .

$$S = H * X \quad (3.11)$$

Les syndromes LDPC sont ensuite accumulés [12]. Les bits de source (X_1, X_2, \dots, X_N) sont additionnés (modulo 2) en accord avec la structure du graphe de Tanner du code LDPCA, afin de générer le syndrome $S = [S_1, S_2, \dots, S_M]^T$. Les bits de celui-ci sont ensuite accumulés (modulo 2) afin de produire le syndrome accumulé $a = [a_1, a_2, \dots, a_m]^T$.

Avant d'envoyer le syndrome accumulé au décodeur, nous sélectionnons le positionnement des bits de syndrome à envoyer, selon le débit imposé par le décodeur. L'encodeur stocke tous les bits de syndrome accumulé, délivré par le codeur LDPCA, dans un tampon et les transmet progressivement à la demande du décodeur. A chaque fois qu'il ne réussit pas à décoder, le décodeur demande l'envoi de 57 bits de syndrome supplémentaires.

Pour le cas de la CE, l'existence d'un canal de retour peut être justifiée par ce que le corps humain supporte la communication bidirectionnelle entre un biocapteur implantable et le terminal de contrôle externe [103-104]. Le support de la communication bidirectionnelle permet une fonctionnalité importante, qui comprend une transmission robuste des données, la navigation et le contrôle des capsules. Par exemple, le système biomédical sans fil a été étudié dans [103] et une communication bidirectionnelle supportée entre le capteur et la borne de commande, où la transmission dans chaque direction a eu lieu en mode semi-duplex. Par conséquent, un canal de retour de l'encodeur au décodeur est une solution fiable pour l'application suivie de la capsule endoscopique sans fil.

3.3.3. Décodeur

Dans le décodeur, les premiers bits de coefficients DC sont décodés par le décodeur Golomb-Rice et le décodeur prédictif. Ensuite, les coefficients DC sont utilisés pour générer la SI. Cette SI est utilisée pour décoder les syndromes qui sont envoyés par le codeur SW.

3.3.3.1. Génération de l'information adjacente

Lorsque les coefficients DC ont été décodés, on utilise la BDD pour générer la SI : les coefficients DC sont comparés aux coefficients DC de chaque image dans la BDD. Pour

accélérer la recherche dans cette dernière, on génère les coefficients DC de chaque image hors ligne et on les stocke dans un tableau. La sélection dans la BDD est effectuée par SAD (*sum absolute difference*) (équation 3.12) où la SAD la plus petite donne l'image la plus corrélée.

$$SAD(I) = \sum_{k=1}^M |DC_x - DC_y| \quad (3.12)$$

Où DC_x les coefficients DC d'image source, DC_y les coefficients DC d'image dans la BDD et M est le nombre des coefficients dans l'image (3762 coefficients).

3.3.3.2. Décodage Slepian-Wolf

Lorsque la SI est disponible dans le décodeur, on applique la DCT sur l'image de la SI et on génère les coefficients DC et AC. Ensuite, ces coefficients sont regroupés sur des bandes de fréquences, et convertis en des plans de bits pour être utilisés dans le décodage des syndromes. Avant de commencer le décodage, le modèle Laplacien est proposé pour estimer les erreurs entre la source et la SI. Dans les systèmes DVC classiques [28], les paramètres du modèle Laplacien sont estimés en ligne par la différence entre les deux images clé compensés en mouvement. Cependant, notre système DVC utilise seulement la trame WZ. Par conséquent, nous ne pouvons pas estimer les paramètres du modèle Laplacien en ligne, c'est pourquoi nous proposons de mettre le paramètre du modèle Laplacien par défaut à $\alpha = 0.5$.

Le décodeur LDPC corrige les erreurs des bits de SI en utilisant les bits syndrome reçus. Un algorithme de Propagation de Croyance (*Belief-Propagation* BP) est utilisé pour le décodage. L'algorithme BP est une approche de décodage soft, qui passe un rapport de vraisemblance logarithmique (LLR) entre les nœuds de la source et les nœuds du syndrome, dans un graphe.

Dans la pratique, l'algorithme BP est exécuté pour un nombre maximum d'itérations de 100 ou jusqu'à ce que le paquet de syndromes soit décodé. Au début du décodage, l'encodeur envoie les syndromes au décodeur par paquet où la taille de chaque paquet de syndromes est de 57 bits. Le décodeur qui reçoit ces données lance l'opération de décodage en exploitant ces bits reçus, l'information adjacente (SI) et le modèle de bruit de corrélation (modèle Laplacien). Si le décodage du paquet en cours ne réussit pas, le décodeur envoie une information au codeur pour l'informer de la nécessité d'envoyer un autre bloc de bits de syndrome, dès que le décodeur reçoit ce bloc de données, il relance l'opération de décodage à nouveau. Cette opération de

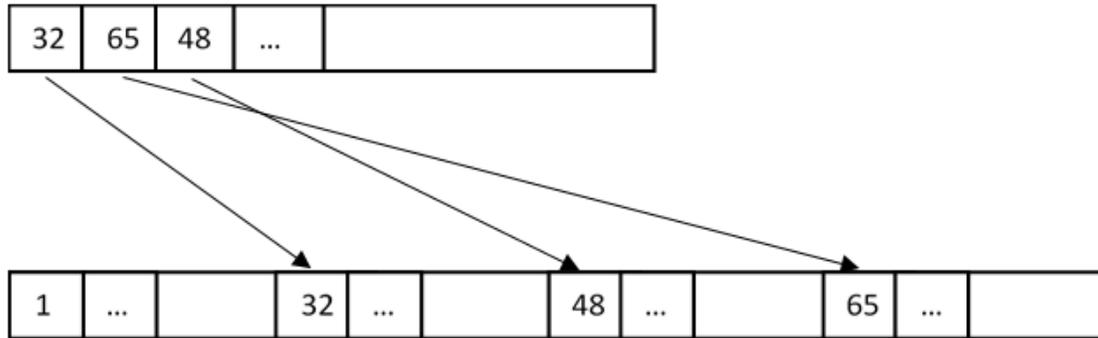


Figure.3.4- Réorganisation des bits de syndrome selon l'ordre de codeur

communication des données se répète itérativement jusqu'à ce que le décodage soit effectué avec succès ou bien jusqu'à l'envoi du paquet complet de syndromes.

Après le décodage de tous les paquets de l'image, le codeur réorganise les bits décodés selon le même ordre du codeur (figure 3.4), et envoie les bits vers le bloc de reconstruction de l'image décodée. Ensuite, le décodeur passe à l'image WZ suivante. Les mêmes opérations précédentes et le protocole de communication sont répétés pour chaque paquet d'images WZ. Dans le cas où le syndrome complet d'un paquet est envoyé au décodeur, et que celui-ci ne puisse pas le décoder, on multiplie le syndrome complet du paquet par l'inverse de la matrice de parité H pour obtenir le paquet décodé.

3.3.3.3. Reconstruction

Lorsque les plans de bits des bandes de fréquences sont tous décodés, nous transformons les bits en des valeurs de coefficients DCT décodées et quantifiées de la trame WZ. Après, la transformation en cosinus discret inverse de H.264 est appliquée pour trouver les pixels décodés. Cette transformation est la même que dans [34] :

$$\hat{X}_{2k}^{BF} = C_{inv} \hat{X}_{2k}^{DCT2} C_{inv}^T \quad (3.13)$$

Où C_{inv} est la matrice de transformation inverse :

$$C_{inv} = \begin{bmatrix} 1 & 1 & 1 & \frac{1}{2} \\ 1 & \frac{1}{2} & -1 & -1 \\ 1 & -\frac{1}{2} & -1 & 1 \\ 1 & -1 & 1 & -\frac{1}{2} \end{bmatrix} \quad (3.14)$$

Pour \hat{X}_{2k}^{DCT2} est les valeurs des coefficients transformés et quantifie calculé comme dans [30].

3.4. Utilisation d'image sous-échantillonnée comme une information de hachage

Dans cette section, nous proposons une autre technique de hachage pour aider à trouver une bonne SI et en même temps on propose de coder les composantes couleurs des images CE. Le système proposé (DVC-HSE) [88] utilise une image sous-échantillonnée comme une information de hachage. Le système DVC-HSE décode en premier la composante luminance, par la SI générée. Pour les composantes couleurs (chrominance), les coefficients DC sont codés par le codage prédictif, ensuite, les coefficients AC des composantes couleurs sont décodés en utilisant la SI. La figure 3.5 représente le système DVC de [88] qui utilise l'image sous-échantillonnée pour générer l'information adjacente (SI).

3.4.1. Encodeur

Dans la section 3.3, nous avons proposé une nouvelle architecture du DVC (DVC-HDC) [87] pour la CE en fonction de la similitude entre les anatomies humaines. Ainsi, nous avons proposé de construire une BDD contenant des images pouvant être prises à partir d'un tractus digestif humain normal ; et l'utiliser comme SI pour décoder d'autres images codées WZ. Dans cette section, nous proposons de coder des composantes couleurs des images CE. Le système proposé (DVC-HSE) [88] exploite les images sélectionnées à partir de la BDD par l'information de hachage pour créer la SI. La manière de la sélection dans BDD est différente de celle de section 3.3, le système utilise une image sous-échantillonnée comme une information de hachage. Le système DVC-HSE décode la composante luminance premièrement par la SI générée, en utilisant l'information de hachage (image sous-échantillonnée) et les images choisies à partir de la BDD. Pour les composantes couleurs (chrominance), les coefficients DC sont codés

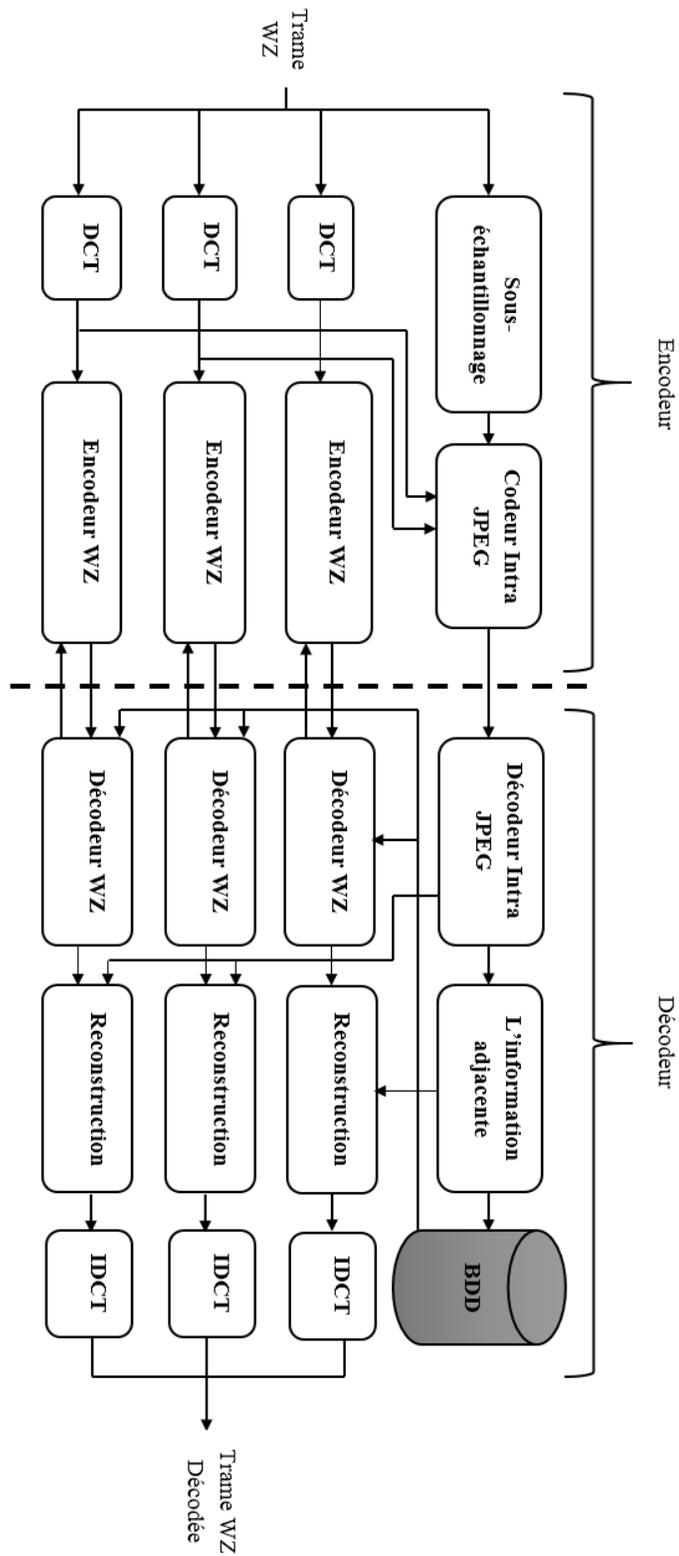


Figure 3.5-Architecture du DVC proposé dans [88].

Chapitre 3 : codage vidéo distribué basé sur une base de données

par le codage prédictif, ensuite, les coefficients AC des composantes couleurs sont décodées en utilisant la SI. La figure 3.5 représente le système du DVC [88] qui utilise l'image sous-échantillonnée pour générer l'information adjacente (SI).

Codage de la composante luminance

Le système DVC-HSE code la composante de luminance en deux parties. Dans la première partie, nous créons une information de hachage, en sous-échantillonnant la trame WZ, la taille de l'image sous-échantillonnée est 128x128 pixels. Ensuite, nous l'encodons avec un intra-code conventionnel (comme JPEG), à une qualité très basse, puis on la transmet au décodeur. Dans la deuxième partie, nous divisons la trame WZ en blocs 4x4 et nous appliquons la DCT. Ensuite, nous organisons les coefficients DC et AC en bandes de fréquences et les quantifions avec les matrices de quantification figure 1.12. Ainsi, nous créons les plans de bits et nous les codons avec le code LDPCA [12]; nous générons les bits de syndrome et nous les envoyons au décodeur.

Codage de la composante couleur

Pour les composantes de chrominance, nous appliquons le DCT 4x4, et nous organisons les coefficients en bandes. Nous séparons ensuite la bande DC des bandes AC. La figure 3.6 explique comment la bande DC est codée. Tout d'abord, nous appliquons le codage prédictif sur les coefficients DC, puis les valeurs résultantes sont encodées par le codage Huffman en utilisant le tableau DC du Huffman utilisé dans le JPEG [105], et on transmet le résultat au décodeur. Les bandes AC des composantes couleurs d'une longueur de 990 valeurs sont codées comme suit : nous quantifions les valeurs par les matrices de quantification proposées, présentées dans la figure 3.7, puis nous convertissons ces valeurs en flux de bits. Nous prenons ces symboles et formons un bloc de bits qui sont envoyés au codeur LDPCA. Les bits de syndrome résultant sont stockés dans un tampon et on transmet un sous-ensemble de ces bits au décodeur, sur demande.

3.4.2. Décodeur

Le décodeur intra-classique (JPEG) décode l'information de hachage et l'utilise pour chercher dans la BDD et pour choisir la meilleure image la plus corrélée à l'image source. La recherche

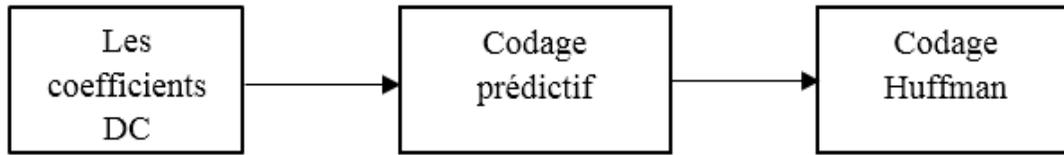


Figure 3.6- Diagramme de codage des coefficients DC

DC	8	0	0	DC	4	0	0	DC	16	8	0	DC	64	16	0
8	0	0	0	4	0	0	0	16	8	0	0	64	16	0	0
0	0	0	0	0	0	0	0	8	0	0	0	16	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Figure 3.7- Matrices de la quantification pour les composantes couleurs

dans la BDD, en utilisant l'équation (3.15), est faite entre l'information de hachage et l'image de la BDD. Ensuite, on choisit l'image qui permet d'obtenir la SAD la plus petite.

$$SAD = |HI(x, y) - ImBDD(x, y)| \quad (3.15)$$

Ou, $HI(x, y)$ et $ImBDD(x, y)$ sont l'information de hachage et l'image sélectionnée respectivement à partir de la BDD. Pour réduire la complexité de la recherche, on met des images de même taille que l'information de hachage dans la BDD.

Ensuite, on combine l'information de hachage et l'image choisie dans la BDD pour générer la SI. Lorsqu'on choisit la meilleure image dans la BDD, on prend l'image correspondante ayant la taille originale, pour l'utiliser dans la génération de la SI. La génération de la SI finale est effectuée entre l'information de hachage et les 5 meilleures images sélectionnées à partir de la BDD. Ceci est nécessaire parce que l'image de hachage peut être très similaire à une certaine image dans une partie et différente dans une autre partie de l'image. Cette opération utilise des blocs 16x16 de l'information de hachage et la compare aux blocs des 5 images de la BDD en utilisant la SAD pour choisir le bloc le plus corrélé.

$$SAD = \sum_{x=1}^{16} \sum_{y=1}^{16} |Bhs(x, y) - Bfr(x, y)| \quad (3.16)$$

Chapitre 3 : codage vidéo distribué basé sur une base de données

Où $Bhs(x, y)$ sont les blocs 16x16 de l'information de hachage et $Bfr(x, y)$ sont les blocs 16x16 des images sélectionnées dans le BDD. La sélection des meilleurs blocs de SI est effectuée selon la règle :

$$\left\{ \begin{array}{l} \text{if}(SAD < T) \\ \text{blocSI} = \text{blocDB} \\ \text{else} \\ \text{blocSI} = \text{blocHI} \end{array} \right.$$

Après plusieurs tests, nous avons fixé le seuil à $T=600$, qui joue un rôle important dans la génération de la SI dans notre système. Lorsque, la SI est générée, nous l'utilisons pour décoder les coefficients DC par notre décodeur WZ. Quand, la première bande est décodée, nous reconstruisons la SI, et nous l'améliorons avec les images de la BDD. Nous répétons le processus d'amélioration de la SI après décodage de chaque bande.

Pour les coefficients AC des composantes chrominances, après avoir sélectionné l'image la plus corrélée avec la source dans la BDD, nous lui appliquons la DCT sur les composantes chrominances (U et V) et la quantifions avec les mêmes matrices pour créer les plans de bits de chaque bande AC. Le sous-ensemble des bits syndromes reçu sert à décoder le flux de symboles. Si le décodeur ne peut pas décoder de manière fiable un symbole, il demande des bits supplémentaires du syndrome à travers le canal de retour. Afin de modéliser la répartition des erreurs entre les bandes AC correspondantes de la trame SI et WZ, nous utilisons une distribution Laplacienne et fixons le paramètre α à 0.5.

La reconstruction effectue l'inverse de la quantification et décode les coefficients AC de la trame WZ. Ensuite, la transformation inverse DCT est effectuée, et la trame entière est restaurée dans le domaine des pixels.

Le codage d'une trame peut se résumer comme suit :

Étape 1 : l'encodeur crée l'information de hachage par sous-échantillonnage de la trame WZ et la code par un codage intra (JPEG).

Étape 2 : l'information de hachage est décodée par un décodeur intra et utilisée pour effectuer une recherche dans la BDD.

Étape 3 : L'information adjacente est générée par l'information de hachage et les images sélectionnées dans la BDD sont utilisées pour le décodage de la composante luminance.

Étape 4 : nous appliquons le DCT sur les composantes couleur et créons les coefficients DC et AC.

Étape 5 : nous codons les coefficients DC par le codage prédictif.

Étape 6 : nous utilisons les images sélectionnées pour décoder les coefficients AC de chrominance par un décodeur WZ.

Étape 7 : nous recueillons les coefficients DC et AC décodés et reconstruisons les trames de chrominance.

Étape 8 : nous combinons la luminance et la chrominance pour reconstruire la trame WZ.

3.5. Résultats expérimentaux

Dans cette section nous allons évaluer l'efficacité de codage et la complexité du codec proposé, par rapport à d'autres codecs vidéo tels que : Motion JPEG [106] et intra H.264/AVC [34]. Nous comparons DVC-HDC avec le codec traditionnel Motion JPEG et intra H.264/AVC. Les différences sont évaluées en utilisant des séquences vidéo de CE représentant diverses zones du tractus gastro-intestinal. Ces séquences contiennent 50 images de la capsule *PillCam SB2*, avec une vitesse d'acquisition de 2 images/seconde et une résolution de 256 x 256 pixels. Motion JPEG (MJPEG) a été utilisé comme référence car cette technique est couramment utilisée dans la capsule endoscopique comme *Pillcam SB2*. Les résultats sont comparés aussi avec intra AVC H264. Dans nos tests, uniquement la luminance est codée dans le système DVC-HDC [87] mais pour le système DVC-HSE [88] toutes les composantes sont codées.

3.5.1. La qualité de l'information adjacente

La qualité de la SI joue un rôle important dans la performance du système de DVC, lorsque la SI est bonne le nombre d'erreurs entre la SI et l'image originale sont petites et le système n'aura pas besoin de beaucoup de bits pour les corriger.

Dans la figure 3.8, on a présenté la qualité de la SI générée par le système DVC-HDC pour deux séquences « capsule vidéo test 1 » et « capsule vidéo test 2 ». Le système DVC-HDC utilise des coefficients DC comme une information de hachage et une BDD contenant 4483 images. Les

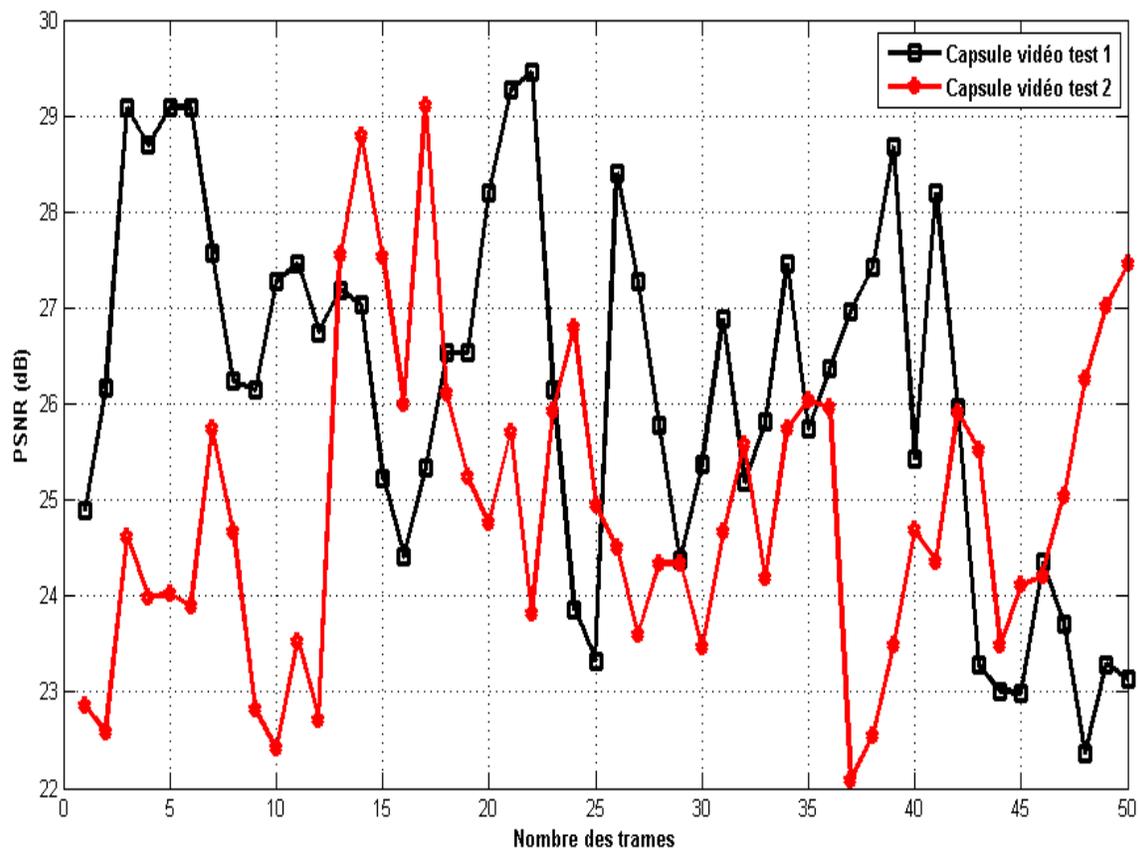


Figure 3.8-Qualité de l'information adjacente de la capsule vidéo test 1 et capsule vidéo test 2 par une BDD contient 4483 images.

deux séquences vidéo de la capsule endoscopique contiennent 50 trames, chaque trame de séquence est comparée avec leur information adjacente par le PSNR.

La variation de la qualité de la SI de la première séquence « capsule vidéo test 1 », est située entre 22.36dB et 29.44dB et donc similaire à la variation de qualité de la SI de la deuxième séquence de « capsule vidéo test 2 », qui varie entre 22.08 dB et 29.10 dB. Par contre, la plupart du temps la qualité de la SI dans la première séquence est supérieure à celle de la deuxième séquence parce qu'elle contient des images simples et claires alors que la deuxième séquence contient plusieurs images complexes. Dans la première séquence, la qualité de la SI peut atteindre 29.44 dB comparée à la deuxième qui ne peut atteindre que 29.10 dB. La qualité moyenne de la première séquence est 26.16dB alors que la deuxième séquence est 24.89dB. Alors, nous pouvons dire que la nature des images dans les séquences de la CE affecte la qualité de la SI.

Chapitre 3 : codage vidéo distribué basé sur une base de données

Généralement, la taille de la BDD affecte directement la qualité de SI. Pour voir ça, la figure 3.9 montre les résultats d'une expérience sur « capsule vidéo test 2 » et l'effet de la taille de la BDD sur la qualité de la SI générée. Pour la première BDD, qui contient 4483 images, la qualité de la SI a atteint 29.10dB, mais avec une BDD qui contient 5988 images la qualité de la SI peut atteindre 29.30dB. La moyenne de la qualité de la SI dans la première BDD est de 24.89dB contre une moyenne de qualité de la SI de 24.98dB, dans la deuxième BDD. En plus, on peut voir que la qualité de la SI générée par la deuxième BDD est meilleure seulement dans quelques trames 24, 45 et 50. Alors, qu'on peut voir clairement lorsqu'on augmente la taille de la BDD la possibilité de trouver une meilleure image, pour l'utiliser comme SI, est grande. Par contre, ceci va introduire plus de complexité au décodeur parce que le temps de recherche dans la BDD va augmenter aussi.

La figure 3.10 montre une comparaison entre la qualité de l'image sélectionnée à partir de la BDD en utilisant les coefficients DC et celle choisie par un hash formé par une image sous-échantillonnée (sans combinaison de 5 images). La figure 3.10 (a) montre une comparaison entre

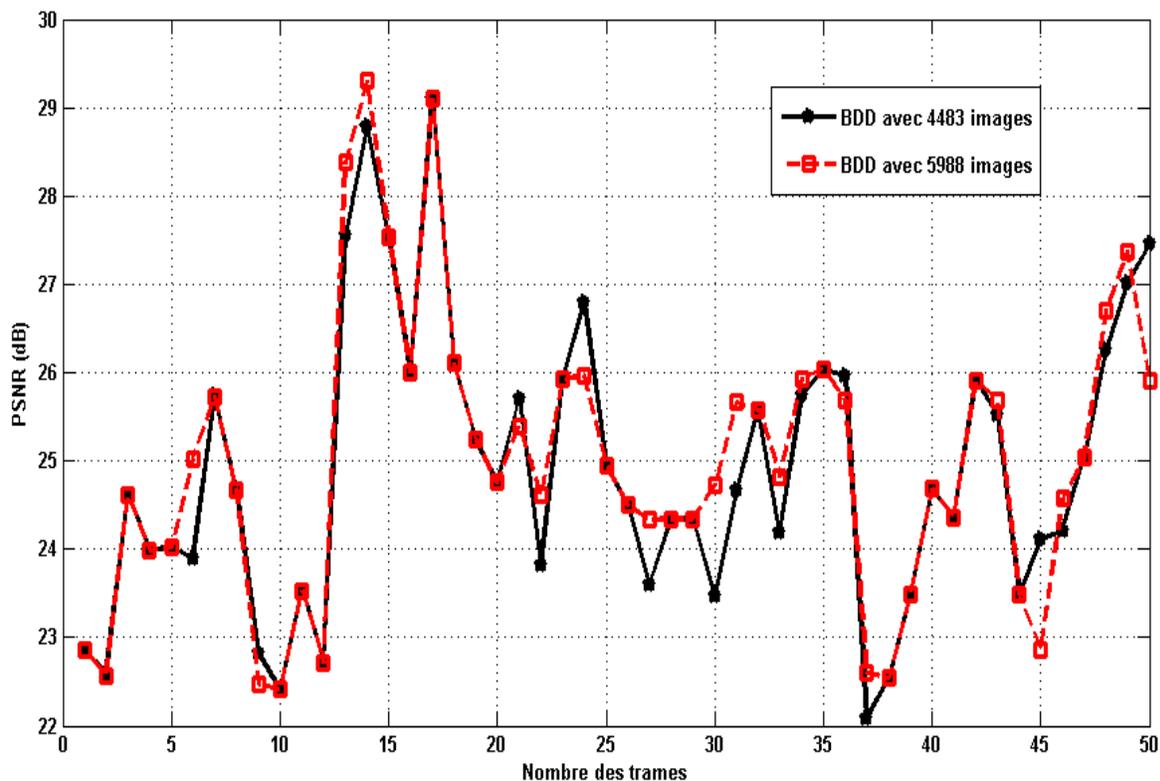
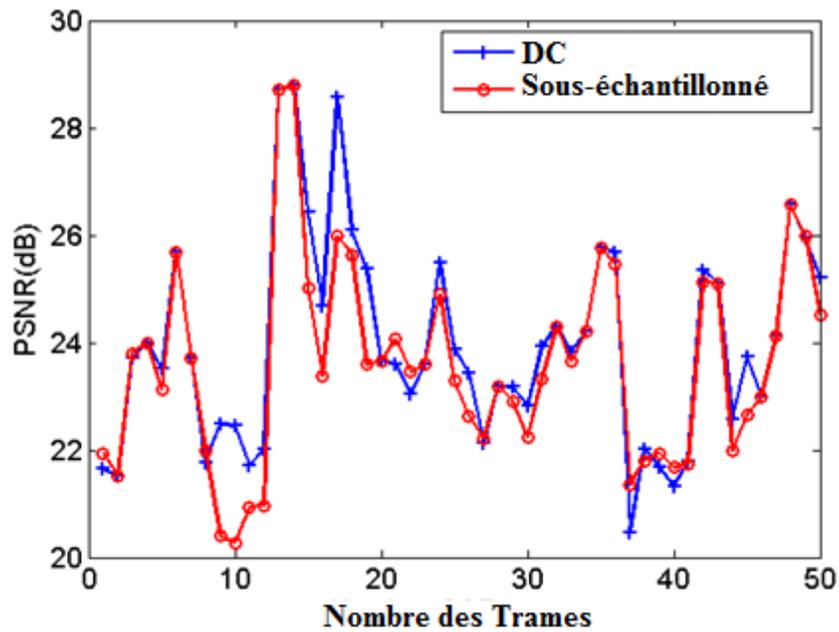


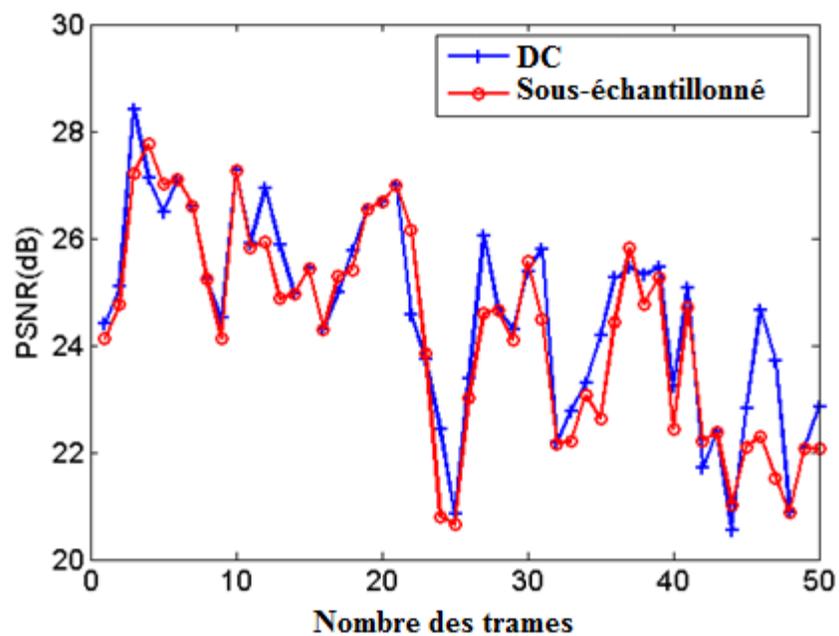
Figure 3.9- Qualité de l'information adjacente de la capsule vidéo test 2 générée par deux BDD.

Chapitre 3 : codage vidéo distribué basé sur une base de données

les deux systèmes DVC-HDC et DVC-HSE pour la séquence de test « Capsule vidéo test 1 ». On peut voir dans cette figure que la sélection de l'image à partir de la BDD en utilisant les



(a)



(b)

Figure 3.10- Qualité d'image sélectionnée à partir de la BDD (Coefficient DC vs sous-échantillonnage)

Chapitre 3 : codage vidéo distribué basé sur une base de données

coefficients DC est meilleure dans 23 images par contre les résultats sont égaux dans 22 images. Donc pour cette séquence, la sélection des images à partir de la BDD en utilisant les coefficients DC est meilleure que la sélection par un hash formé par une image sous-échantillonnée seulement.

Pour la deuxième séquence « Capsule vidéo test 2 », la figure 3.10 (b) montre que la sélection de l'image à partir de la BDD en utilisant les coefficients DC est meilleure dans 25 images et la sélection de l'image à partir de la BDD en utilisant image sous échantillonnée est meilleure dans 8 images. Donc, la sélection des images à partir de la BDD en utilisant les coefficients DC est meilleure même pour la deuxième séquence.

Alors, on peut conclure que la sélection d'une image de la CE à partir de la BDD en utilisant les coefficients DC est meilleure par rapport l'image sous-échantillonnée.

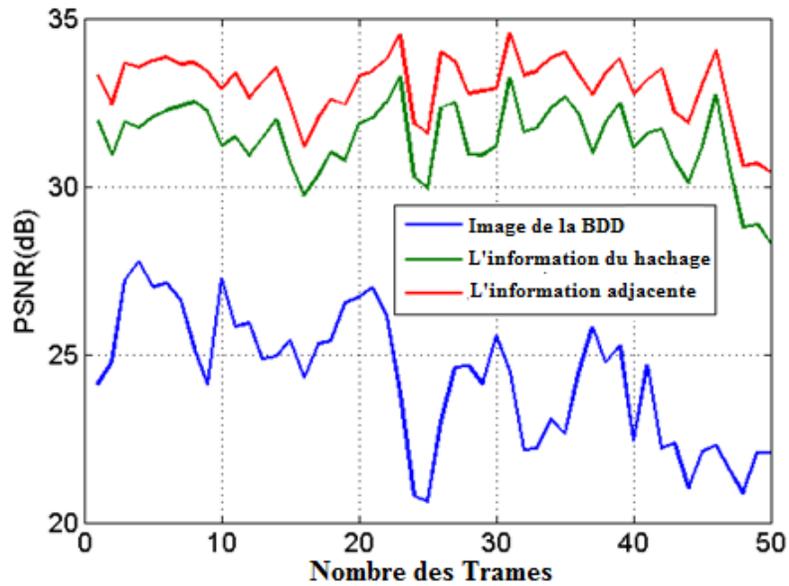
La figure.3.11 présente la qualité de la SI générée par le système DVC-HSE en utilisant la BDD qui contient 4483 images de CE. La figure présente aussi la qualité de l'information de hachage et l'image sélectionnée à partir de la BDD. On peut voir clairement que la qualité de toutes les images de SI est supérieure à 30 dB. La Figure 3.11 (a) montre le PSNR obtenu pour la première meilleure image dans BDD, l'information de hachage et SI pour les 50 images de la séquence. La qualité d'image sélectionnée à partir de la BDD peut atteindre 28.5dB. La figure 3.11 (a) montre aussi que la qualité de la SI peut atteindre 34.7 dB dans l'image 31 et la qualité minimale se trouve dans l'image 50 avec 30.6 dB.

Selon les résultats obtenus, on peut voir que la qualité de la SI est très élevée par rapport à la qualité de l'information de hachage et la meilleure image choisie à partir de la BDD. Par exemple, la qualité de la SI a été améliorée de plus de 2dB par rapport à la qualité de l'information de hachage pour l'image 44 de la séquence. Alors, on peut déduire que la technique de la génération de la SI en utilisant la BDD et l'information de hachage a été efficace et donne de très bons résultats.

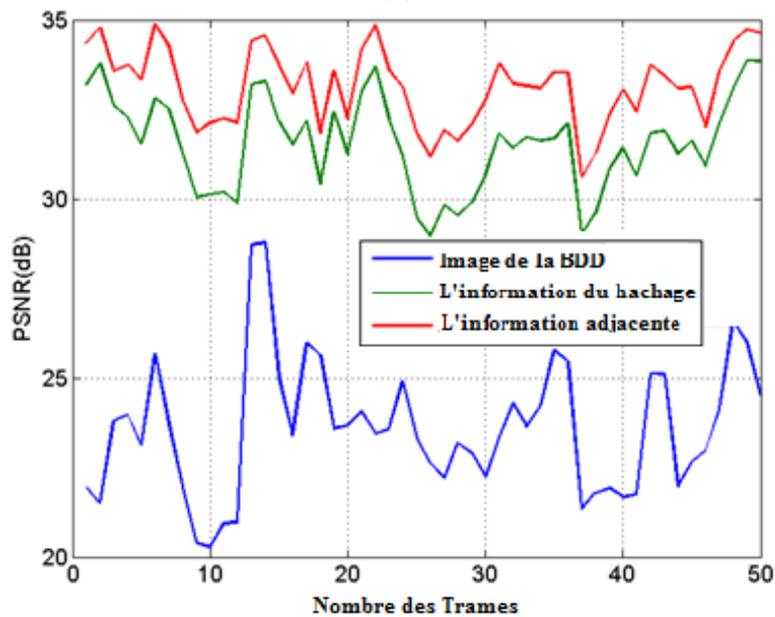
La Figure 3.11 (b) montre le PSNR obtenu pour la meilleure image sélectionnée à partir de la BDD, l'information de hachage et la SI pour 50 images de la deuxième séquence. La figure montre que la qualité de la SI peut atteindre 34.9 dB pour les images 6 et 22 et une qualité minimale de 28.68 dB pour l'image 36. Donc, la qualité de la SI a été améliorée de plus de 2dB par rapport à l'information de hachage. Selon les résultats obtenus, on peut voir que la qualité de la SI est très élevée par rapport à la qualité de l'information de hachage et l'image de la BDD.

Chapitre 3 : codage vidéo distribué basé sur une base de données

Alors, on peut voir qu'il y a une influence directe de la qualité de l'image de la BDD et l'information de hachage sur la qualité de SI générée.



(a)

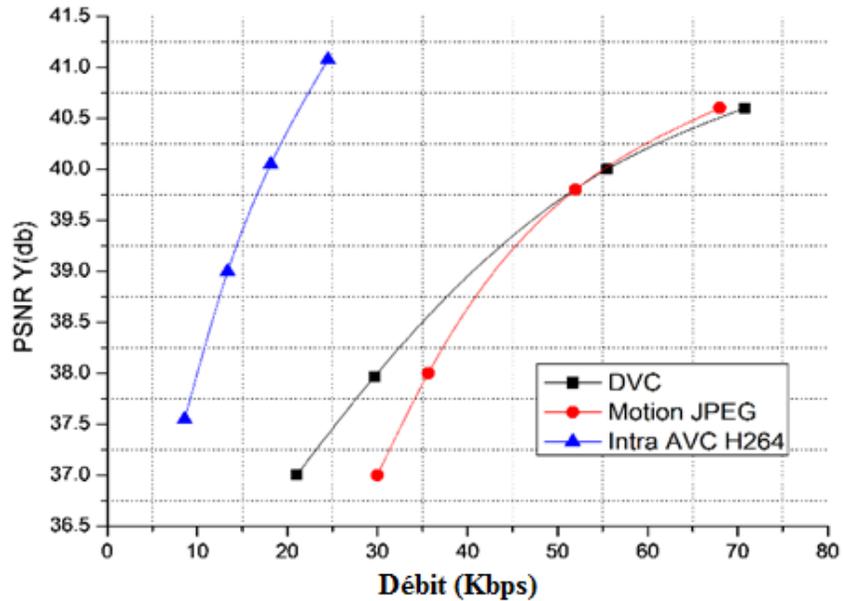


(b)

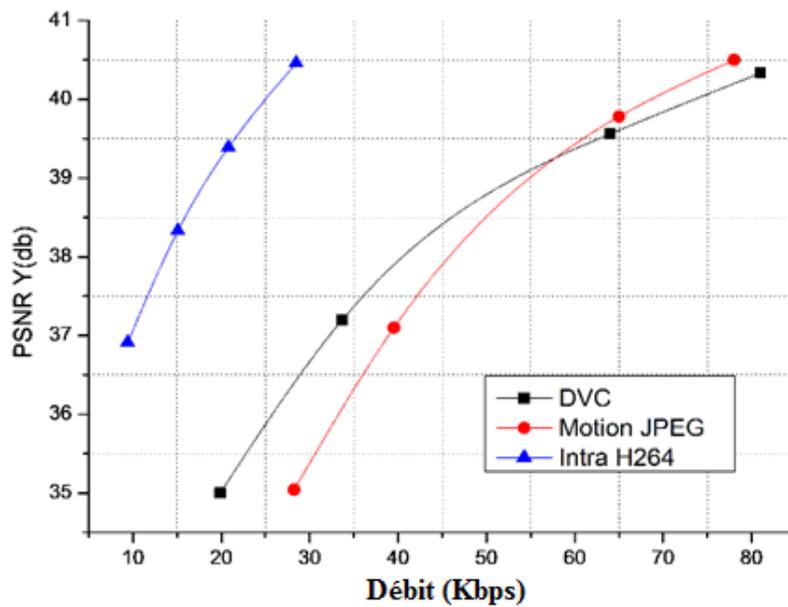
Figure 3.11- Qualité de l'information adjacente générée en utilisant l'information de hachage et l'image de la BDD.

3.5.2. La performance de codage du système

La figure 3.12 montre les quatre points RD (rate-distorsion) de MJPEG et 4 points correspondants aux points 28, 30, 32 et 35 du paramètre de quantification QP de H.264 intra et



(a)



(b)

Figure 3.12-Résultats expérimentaux sur les capsules endoscopiques : (a) "Capsule Test Vidéo 1" et (b) "Capsule Test Vidéo 2".

Chapitre 3 : codage vidéo distribué basé sur une base de données

aussi quatre points RD correspondant aux matrices de quantification (QM) 0, 1, 6 et 7 dans [28], pour le système proposé DVC-HDC. La figure 3.12 (a) montre la comparaison des performances du système DVC [87] avec intra H.264 et Motion JPEG. L'axe X montre le débit binaire luminance, tandis que l'axe des Y montre le PSNR (dB) de la composante luminance. La figure 3.12 montre que le codec H.264 surpasse Motion JPEG et le DVC proposé dans les deux cas, mais le H.264 [33] est un codec très complexe à mettre en œuvre dans les capsules endoscopiques.

Pour les paramètres de quantification les plus élevés (compression haute), le PSNR obtenu par DVC est meilleur (jusqu'à 1dB) par rapport à celui de MJPEG. En outre, pour les paramètres de quantification les plus faibles (compression faible), le PSNR obtenu avec DVC est inférieur à 0,15 dB par rapport à celui de MJPEG.

La figure 3.12 (b) montre la performance RD du système DVC [87] dans la deuxième séquence de la CE. Pour les débits faible (grande quantification et grande compression), le PSNR obtenu par DVC est meilleur jusqu'à 1.5 dB que celui de MJPEG. Par contre, pour les débits élevés, le PSNR de DVC est inférieur par 0,6 dB à celui de MJPEG.

Pour mieux exprimer la différence entre le codec DVC proposé et MJPEG, nous avons évalué la performance de codage en termes de la métrique Bjontegaard Delta (BD) [108]. En particulier, dans "Capsule Test Vidéo 1" et "Capsule Test Vidéo 2", le codec proposé est plus performant que MJPEG et apporte des BD de 22.8721% et 20.3596%, respectivement. En général, on peut remarquer que le codec DVC proposé est toujours supérieur à MJPEG lorsqu'on a une forte compression, contrairement au cas d'une faible compression.

Dans la figure 3.13, une comparaison visuelle est faite entre les images décodées par le DVC-HDC et les images originales de la séquence de la CE. Le codec DVC proposé produit une qualité visuelle qui ne souffre pas d'artefacts de blocs. En fait, dans la figure 3.13, on ne peut voir aucune différence entre l'image décodée par le codec DVC-HDC et l'image originale. Cette qualité confirme la possibilité d'utiliser le DVC dans les applications d'imagerie médicale, où la qualité visuelle élevée est une demande fondamentale.

La performance RD du système [88] est présentée dans la figure 3.14 où seule la composante de luminance est codée. Quatre points de quantification sont dessinés correspondant aux matrices de quantification 1, 2, 3, et 4 de la figure 1.12. La figure 3.14 montre aussi une comparaison entre le système DVC et MJPEG. La figure 3.14 (a) montre la performance RD pour la séquence

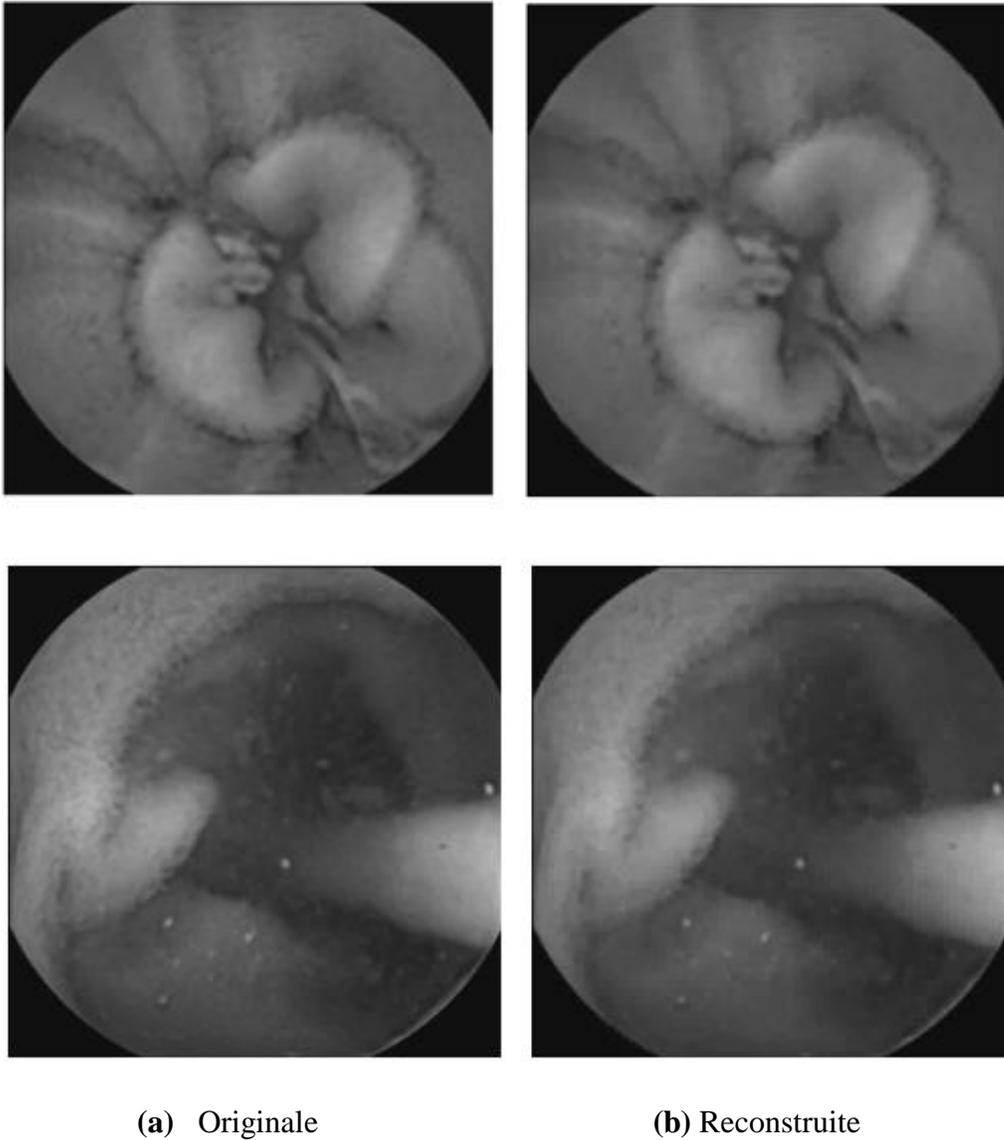


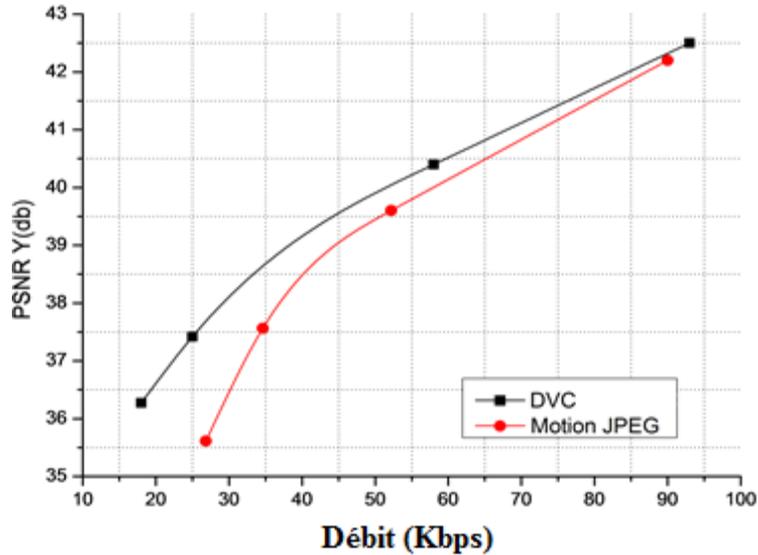
Figure 3.13-Qualité visuelle à partir de la séquence décodée "Capsule Test Video1" et "Capsule Test Video2" : (a) image originale, (b) image décodée par le codec proposé.

'capsule vidéo test 1'. On peut voir que le système DVC proposé est plus performant que MJPEG par 2 dB pour le faible débit et 0.2 dB pour les hauts débits. La figure 3.14 (b) montre la performance RD de la séquence 'capsule vidéo test 2'. Pour cette séquence, le système DVC est aussi plus performant que MJPEG avec un gain de 1 dB pour les faibles débits, mais pour les hauts débits MJPEG est plus performant (à 0.1 dB).

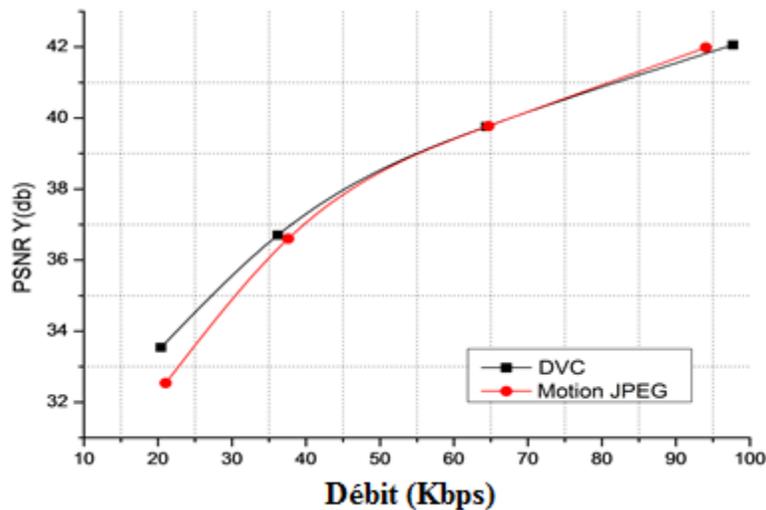
Chapitre 3 : codage vidéo distribué basé sur une base de données

La performance RD des composantes de chrominance est présentée dans la figure 3.15 pour les deux séquences CE de test. Les composantes de chrominance du système DVC sont aussi comparées aux composantes de chrominance codées par MJPEG.

Les résultats sur les composantes de chrominance de la séquence de test vidéo 1 sont présentés sur la figure 3.15 (a). La compression de la composante V de notre système DVC est plus performante lorsqu'on la compare avec la composante V de MJPEG; que ce soit pour une



(a)

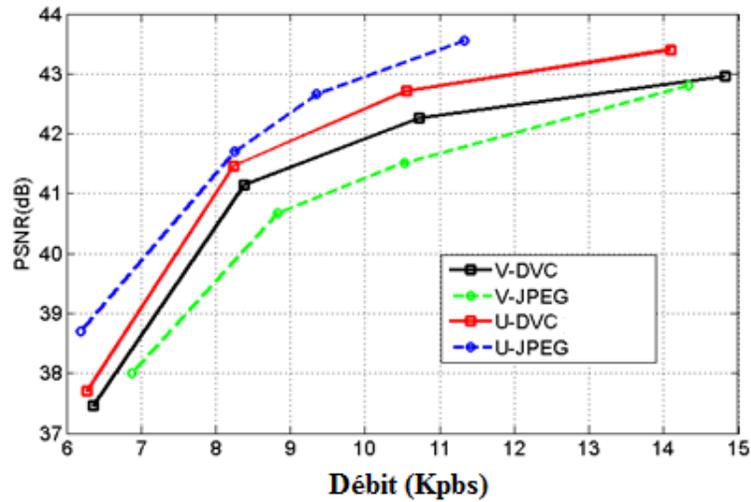


(b)

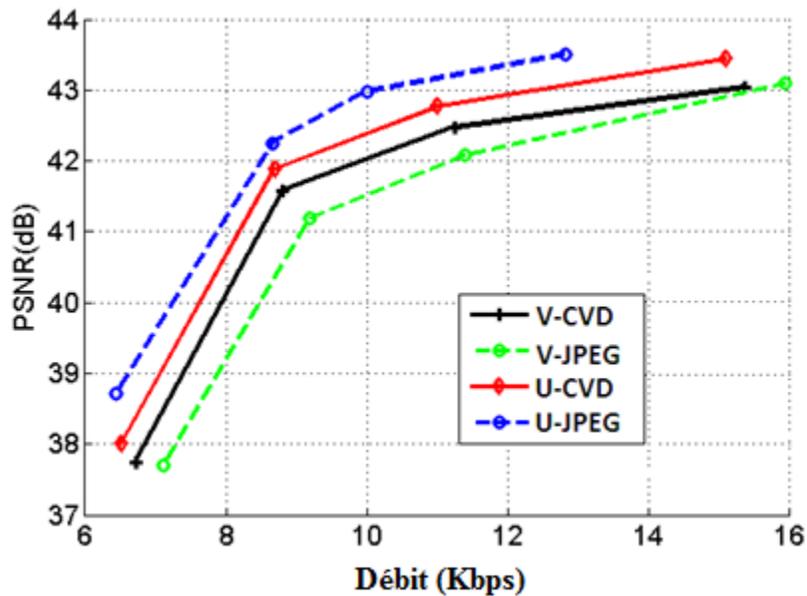
Figure 3.14- Résultats expérimentaux d'une capsule endoscopique sans fil : (a) "Capsule Test Vidéo1" et (b) "Capsule Test Vidéo 2" pour le composant luminance.

Chapitre 3 : codage vidéo distribué basé sur une base de données

compression forte ou faible. Cependant, la compression de la composante U de MJPEG est plus performante lorsqu'on la compare avec la composante U de notre système DVC [88]. La figure 3.15 (b) présente les mêmes résultats pour la séquence vidéo test 2. Dans ce cas aussi, la compression de la composante V du système DVC est meilleure que MJPEG, contrairement à la



(a)



(b)

Figure 3.15-Résultats expérimentaux sur d'une capsule endoscopique sans fil: (a) "Capsule Test Vidéo 1" et (b) "Capsule Test Vidéo 2" en codant les composantes couleurs.

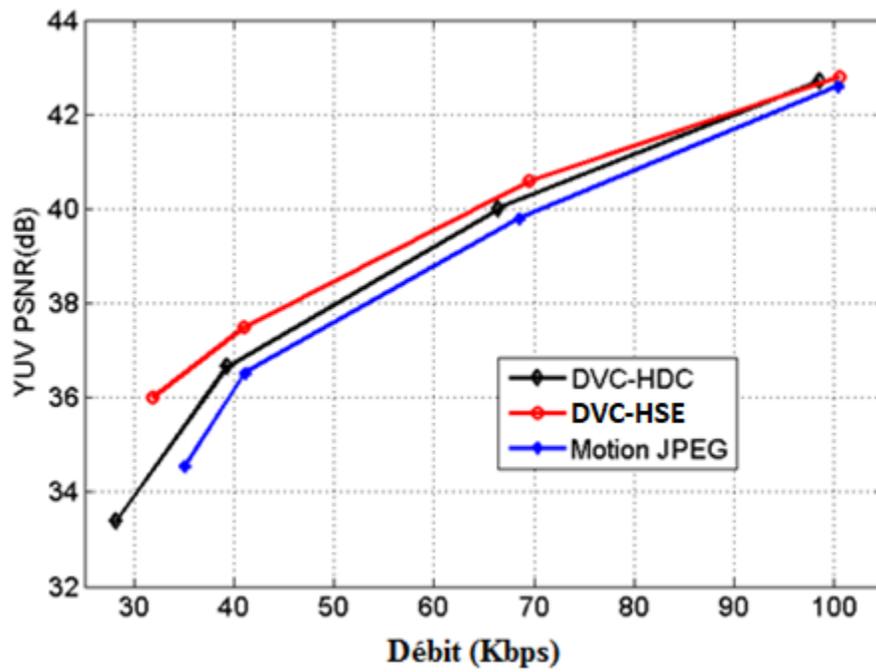
Chapitre 3 : codage vidéo distribué basé sur une base de données

composante U où MJPEG est meilleur. Pour comparer la performance globale, nous avons exprimé la différence de performance entre le codec proposé et MJPEG en termes de la métrique Bjontegaard Delta (BD). Dans ce cas, le codec proposé surpasse clairement MJPEG. Pour "capsule test vidéo 1", le codec proposé apporte des moyennes BD de 4.2487% et la qualité 0.0615 dB contre MJPEG. Pour "capsule test vidéo 2" le gain est de 2.2914% et 0.0934 dB contre MJPEG.

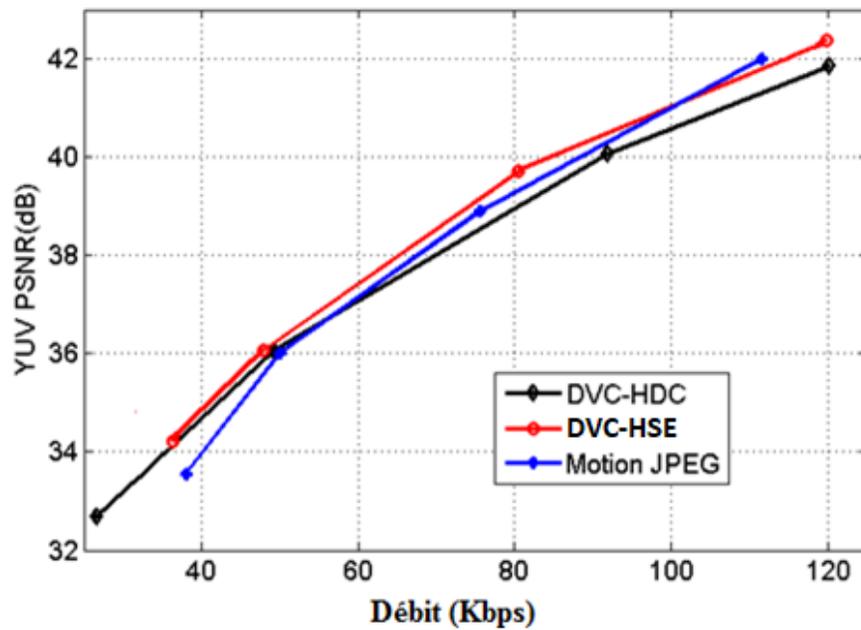
Une comparaison entre les deux systèmes proposés avec le MJPEG, où on code tous les composantes couleurs de l'image, est présentée dans la figure 3.16. Les composantes de chrominance dans le système DVC-HDC utilisent les composantes de chrominance de l'image sélectionnée par la BDD comme SI pour la chrominance. Ces composantes chrominances sont ensuite codées comme dans le système DVC-HSE. La figure 3.16 (a) montre les performances RD de nos systèmes DVC [87,88] en utilisant la séquence "capsule test vidéo 1". La figure 3.16(a) montre que le système DVC-HSE est plus performant que le système DVC-HDC et MJPEG pour une compression forte ou faible, parce que la qualité de SI de système DVC-HSE est meilleure que celle du DVC-HDC. On rappelle que le système DVC-HDC utilise seulement l'image sélectionnée à partir de la BDD comme la SI directement. Cependant, le système dans DVC-HSE, la SI est générée en utilisant l'information de hachage et cinq images sélectionnées à partir de la BDD. En plus, dans le système DVC-HSE, après décodage chaque bande, on améliore la qualité de la SI et cela permet de réduire les erreurs entre la source et la SI et augmente la performance RD du système DVC-HSE.

Pour la deuxième séquence, dans la figure 3.16 (b), le système DVC-HSE est plus performant en le comparant à MJPEG et le système DVC-HDC lorsque la compression est haute. Cependant, lorsque la quantification est faible, le MJPEG est meilleur. Mais en termes de la BD, notre système DVC-HSE est plus performant que MJPEG avec 0.6207% et une qualité de 8.8224 dB.

Enfin, la figure 3.17 présente une comparaison visuelle entre l'image originale et l'image décodée en utilisant le système DVC-HSE proposé. Nous pouvons voir qu'il n'y a pas de différences visibles entre le résultat du système DVC et l'image originale. Le codec proposé produit la même qualité visuelle et ne souffre pas d'artefacts de blocs.



(a)



(b)

Figure 3.16- Résultats expérimentaux sur les capsules endoscopiques : (a) "Capsule Test Vidéo 1" et (b) "Capsule Test Vidéo 2", en codant tous les composants d'image.

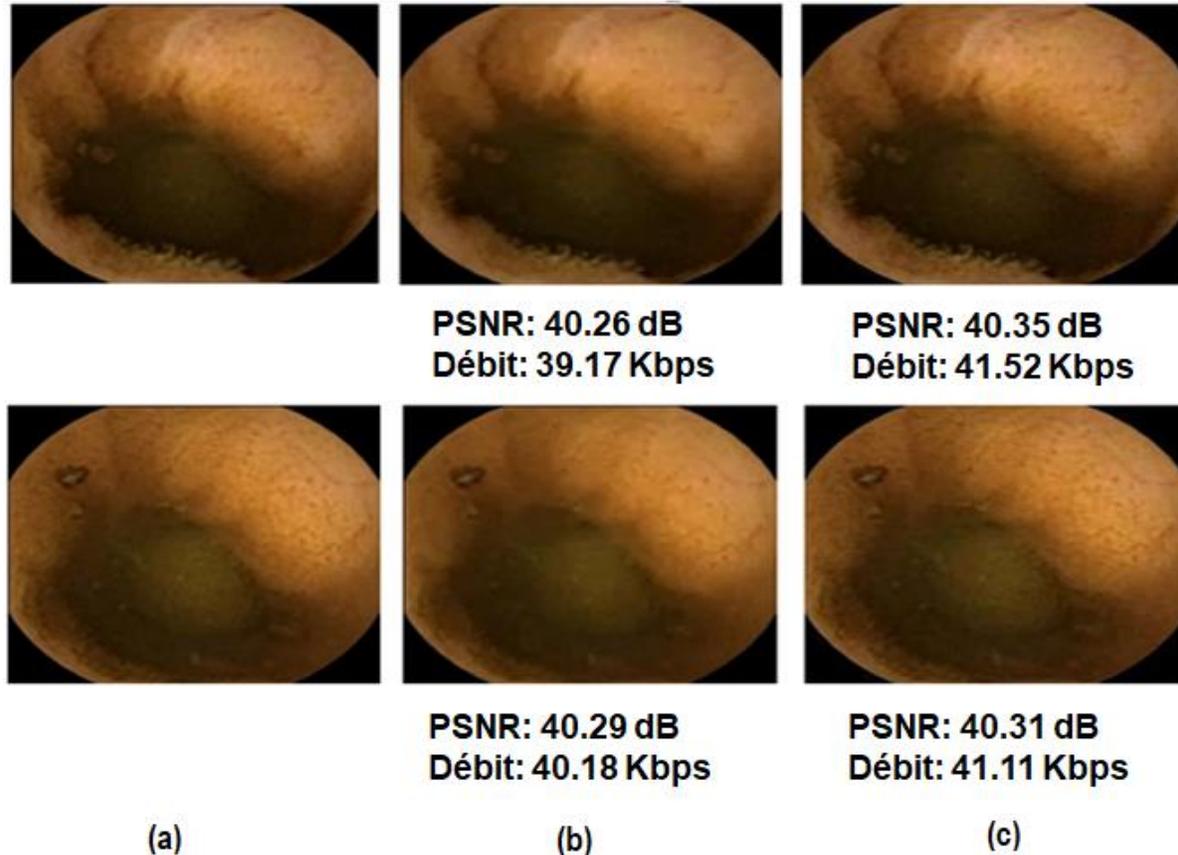


Figure 3.17– La qualité visuelle d’une image décodée par le système DVC [88]. (a) Original, (b) Motion JPEG et (c) Système DVC

3.6. Conclusion

La simplicité de l’encodeur est un facteur important dans les applications de la transmission via un réseau sans fil. Ce qui économise l’espace dans la mémoire et le nombre de calculs utilisé dans l’encodeur, qui utilise uniquement des opérations d’implémentation simples telles que la DCT, quantification, codeur Golomb-Rice et codeur SW. L’approche proposée est innovante par rapport au travail référencé. Une expérimentation détaillée montre que le codec DVC proposé dépasse le MJPEG en débit faibles, tout en conservant une faible complexité d’encodage. Cependant, pour améliorer la performance de codec, pour tous les débits, il faut d’abord améliorer la qualité de la SI. Donc, dans le chapitre 4, nous proposons une nouvelle architecture pour améliorer la génération de l’information adjacente.

Chapitre 4

Codage vidéo distribué basé sur la quantification vectorielle

4.1. Introduction

Dans les premières solutions que nous avons proposées dans le chapitre 3, nous devions augmenter largement le nombre d'image pour avoir une bonne compression. Dans ce chapitre, un nouveau système (DVC-VQ) [109] est proposé pour la CE. C'est une extension des deux précédentes contributions. Ce système DVC utilise la quantification vectorielle (QV) pour générer une meilleure SI et pour contrôler le budget énergétique, par le décodeur. Le codage dans ce cas se fait en deux étapes : une QV simple est utilisée pour produire une bonne SI, suivie d'un DVC scalable qui améliore progressivement la qualité de l'image décodée. Dans ce chapitre on va détailler les différentes opérations de ce système.

4.2. La quantification vectorielle

La quantification scalaire peut rendre les valeurs des pixels dans l'image plus faibles. Ce qui permet une bonne exploitation de la corrélation entre les valeurs quantifiées. Shannon dans sa théorie montre qu'on peut améliorer la compression des données en codant les vecteurs plutôt que les scalaires [110]. La quantification vectorielle est une technique de compression avec perte, développée par Gray [111]. Elle a pris une place très importante dans le domaine de la compression. La QV est utilisée dans plusieurs domaines du traitement du signal comprenant la compression des données et des images. La procédure de la QV pour la compression d'image comprend trois parties principales : la génération d'un dictionnaire, l'encodage et le décodage [112].

4.2.1. Définition

La QV consiste à convertir un vecteur $x = (x_i ; i = 1, \dots, k)$ de dimension k dont les composantes sont des variables aléatoires réelles et continues, en un vecteur $w(w_i ; i = 1, \dots, k)$ de même dimension, dont les composantes sont des variables réelles et discrètes appelées mots de code (CW : *codeword*). Le CW appartient à un ensemble fini et ordonné de représentants de tous les CWs possible ; cet ensemble est appelé dictionnaire (CB : *codebook*). Les éléments du vecteur à coder peuvent être des pixels, des coefficients de transformée, des valeurs d'erreurs de prédiction ou des écarts avec le vecteur représentatif choisi lors d'une première quantification vectorielle.

4.2.2. Codage de la Quantification Vectorielle

En général, pour coder un vecteur x , un représentant $w = Q(x)$ est sélectionné dans le dictionnaire selon la règle du plus proche voisin, et l'indice du vecteur sélectionné est transmis. Dans le codage à QV des images, le codage est effectué en divisant une image en un groupe de vecteurs (blocs) V_i . Pour coder un bloc, l'encodeur recherche dans le CB, le meilleur CW qui permet de représenter ce bloc avec la plus petite distorsion (figure 4.1). L'encodeur transmet l'indice du CW seulement au décodeur. Ainsi, l'encodeur est très complexe, car le processus de la recherche et le calcul de distorsion sont très complexes.

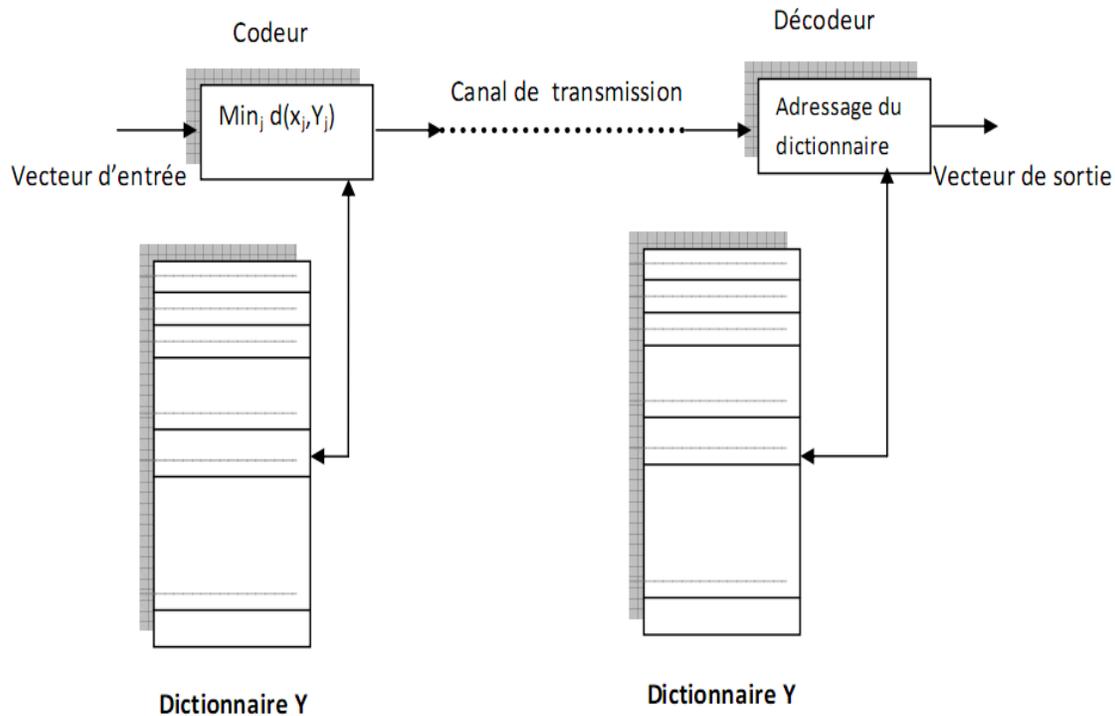


Figure 4.1-Conception classique d'un système de la quantification vectorielle.

4.2.3. Décodage de la QV

Le décodeur utilise uniquement les indices, pour sélectionner les CWs dans le CB et les utilise pour reconstruire l'image décodée. Donc, le décodeur est très simple, car il utilise uniquement les indices reçus pour lire directement les CWs du CB.

Donc, le fait que l'encodeur de la QV prend beaucoup de temps pour la recherche de l'indice et a besoin d'un espace dans la mémoire pour stocker le CB, fait que VQ est une technique complexe pour la CE sans fil, bien que la QV permette une très bonne performance RD.

4.3. L'architecture DVC proposée

Dans le chapitre 3, on a utilisé une BDD qui contient des images complètes pour la génération de la SI. Malheureusement, pour obtenir une très bonne qualité de la SI, il est nécessaire d'avoir une BDD avec un très grand nombre d'images, afin d'assurer une très bonne correspondance pour chaque image possible. En outre, la comparaison entre les deux images de

Chapitre 4 : codage vidéo distribué basé sur la quantification vectorielle

taille normale prend beaucoup de temps. Par conséquent, le temps de recherche pour obtenir une bonne SI était trop long. En plus, compte tenu de la forte similitude entre les images de la CE dans la texture et le contenu, il serait préférable de les diviser en blocs et de prendre de chaque image uniquement les blocs non redondants pour former la BDD. C'est exactement ce que fait la QV. Cela diminuera la taille de la BDD, qui deviendra maintenant un dictionnaire (CB), et permet donc de trouver une bonne SI pour beaucoup d'images. Logiquement, avec cette technique, il est possible de produire une bonne qualité de SI, si nous acceptons l'utilisation d'un CB de très grande taille. Heureusement, grâce au contenu réduit des scènes de la CE, le CB pour tous les blocs possibles pourrait être de petite taille. Cependant, comme on vu plus haut l'implémentation de la QV dans un système à ressources limitées, tel que les CE, est non pratique. Donc, dans notre système, nous proposons de modifier la conception de la QV pour rendre l'encodeur plus simple, en déplaçant les opérations complexes de recherche vers le décodeur. Cette modification va avoir des influences sur le fonctionnement du système DVC et sur l'adaptation de la QV dans le système, et aussi sur les étapes qui conduisent à la génération de la SI et enfin au codage SW.

L'encodeur de la QV classique est très complexe par rapport au décodeur [111] [113], parce que c'est lui qui effectue la recherche dans le CB. Donc, pour l'utiliser avec notre DVC, nous déplaçons les opérations complexes de la recherche et le CB vers le décodeur (figure 4.2.). Par conséquent, toutes les opérations de recherche sont effectuées dans le côté qui contient plus de ressources. Dans ce cas, l'encodeur calcule et envoie une information de hachage et le rôle du décodeur est de trouver dans le CB, le meilleur bloc correspondant à l'information de hachage.

Malheureusement, cette opération réduira les performances de la QV, et certains blocs de l'image reconstruite par la QV peuvent ne pas être de bonne qualité comme ceux de la QV classique, parce que nous n'utilisons pas le bloc d'origine. Dans notre cas, cette réduction de performance peut être acceptable, car l'image de la QV est utilisée uniquement comme SI et elle sera améliorée par les bits de correction d'erreur du codage SW. Pour certains blocs, nous pouvons obtenir le même résultat que celui de la QV classique. En plus ici, l'information de hachage (pas un indice) est transmise, ce qui entraîne une performance RD inférieure. Dans notre cas, il est possible d'utiliser n'importe quel hash pour la QV, mais le compromis de RD sera celui de produire la meilleure qualité de la QV avec le plus petit nombre possible de bits de hachage. Du côté du décodeur, cette

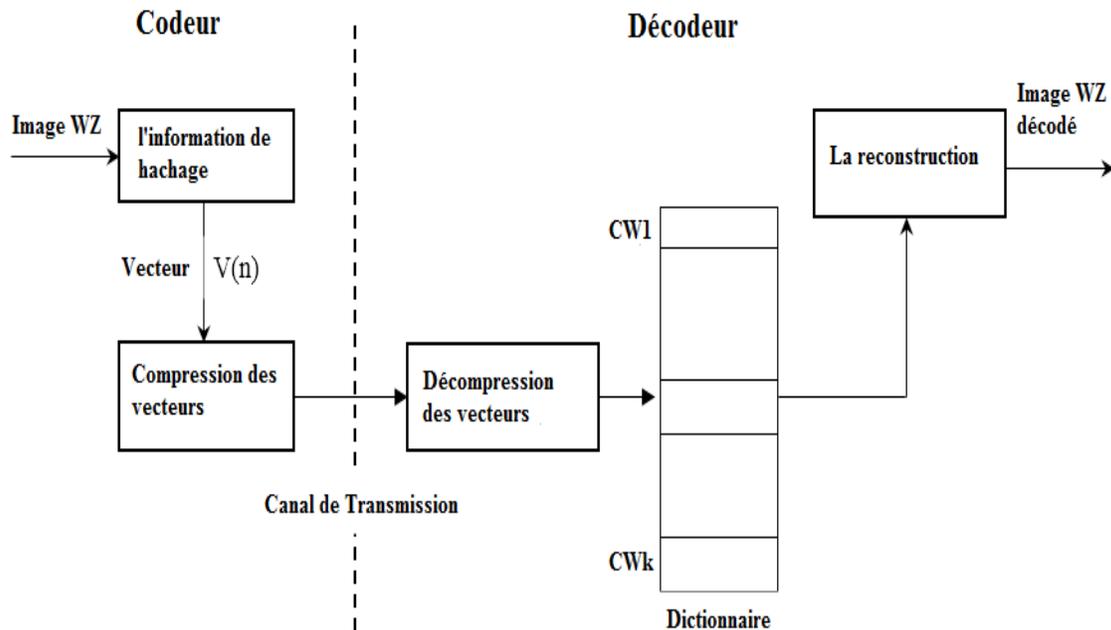


Figure 4.2-La conception de la QV proposée pour le système DVC

performance RD est également affectée par le contenu du CB et les algorithmes de la recherche utilisés pour sélectionner les meilleurs CWs. Une grande taille du CB avec une recherche exhaustive produira une très bonne qualité de SI. Cependant, utiliser un algorithme de recherche rapide et une petite taille du CB peut rendre la recherche plus rapide et le système de la QV simple. Le système DVC complet, que nous avons construit autour de la nouvelle QV, déplace la complexité vers le décodeur. Il est aussi important de noter que le codage QV des trames WZ utilise un seul CB qui est placé côté décodeur. En plus, comme pour nos autres systèmes, présentés précédemment, l'avantage majeur de ce système est qu'il n'a pas besoin de trames clés du tout. Donc, ce système est plus pratique que celui de la figure 1.11, car il n'est pas nécessaire d'implémenter le codage intra dans le codeur.

4.3.1. l'encodeur

Notre encodeur de la QV est simple, il crée un vecteur du hash pour chaque bloc et le comprime en utilisant un codage prédictif et un codage Golomb-Rice. Comme information de hachage, nous avons décidé d'utiliser la valeur moyenne des blocs. Les valeurs moyennes sont très faciles

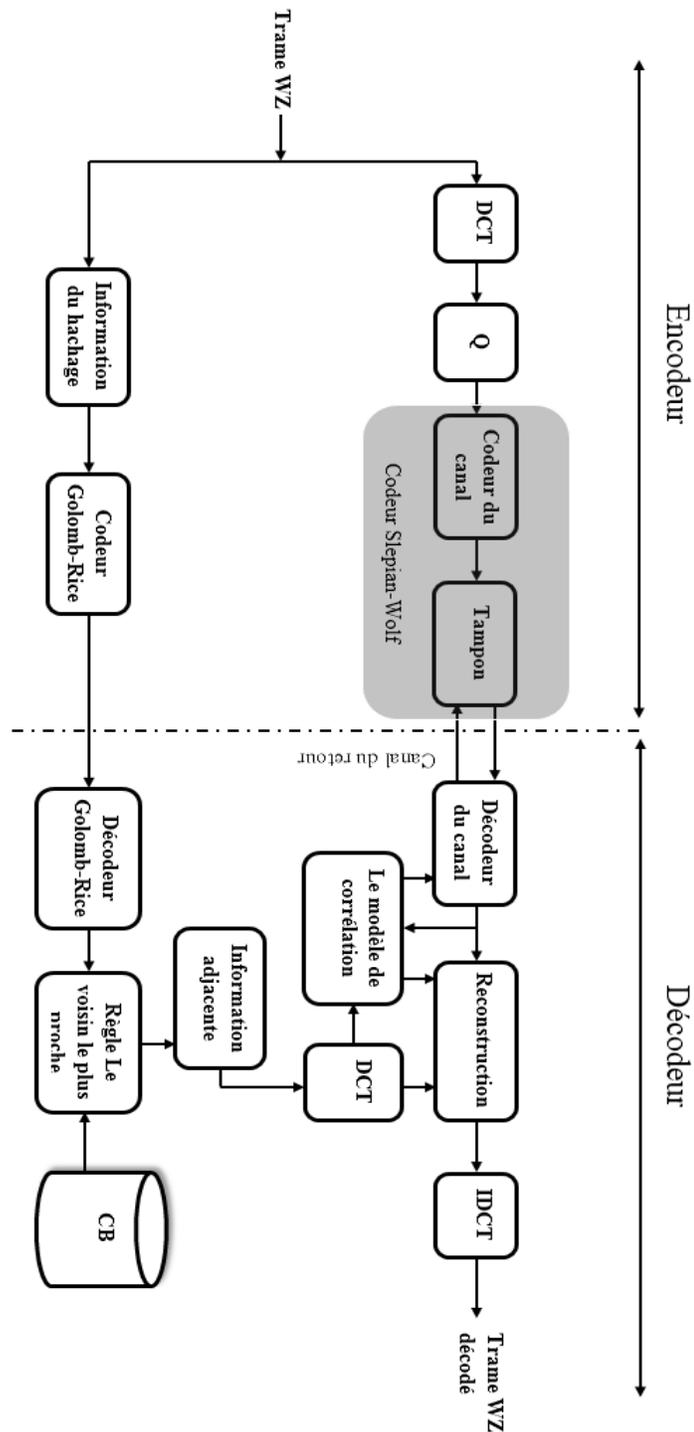


Figure 4.3- Un schéma du système du DVC proposé

Chapitre 4 : codage vidéo distribué basé sur la quantification vectorielle

à calculer et contiennent des informations sur le bloc d'origine qui peuvent être utilisées pour corriger la SI. L'encodeur divise chaque image en macro-blocs de 16x16 pixels (taille du CW). Puis, chaque macro-bloc est divisé en blocs 4x4. Pour chaque bloc 4x4, nous calculons la valeur moyenne. Ensuite, nous regroupons les valeurs moyennes obtenues dans un vecteur, qui représentera l'information de hachage du macro-bloc. En outre, pour réduire davantage le débit de hachage, nous utilisons une combinaison de seulement quatre blocs 4x4 et trois 8x8, au lieu d'utiliser 16 blocs 4x4, comme le montre la figure 4.4. Le codage prédictif est utilisé, entre les valeurs moyennes consécutives, pour réduire davantage leur coût en bits. Enfin, les erreurs de prédiction sont encodées avec le codeur Golomb-Rice comme dans [99] avec un paramètre de la longueur maximale de code $g_{limit} = 23$. La partition utilisée est également transmise avec le hash (2 bits). Par exemple, pour une CE avec une résolution de 256x256 pixels, il existe 256 macroblocs et il n'y a que 1792 valeurs moyennes à coder pour une image. Par rapport à l'utilisation de 16 valeurs par macrobloc, nos tests ont montré que le coût en bits du hash modifié était plus petit et avec seulement une petite dégradation de la qualité de la SI générée. De plus, comme les trames

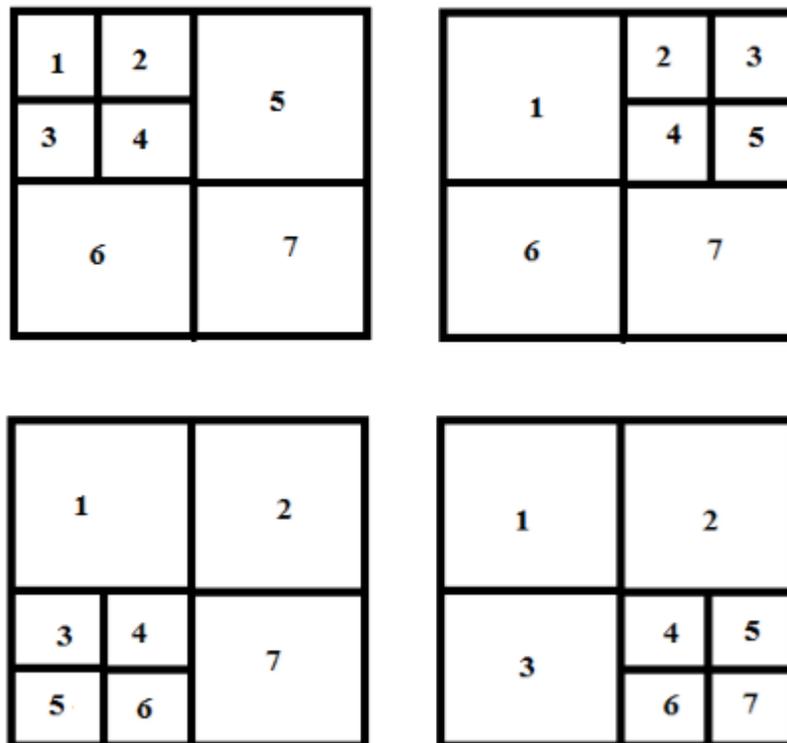


Figure 4.4-Partitions de bloc 16x16

Chapitre 4 : codage vidéo distribué basé sur la quantification vectorielle

de la CE sont représentées dans une zone circulaire, en raison de la forme cylindrique du tractus GI, les coins noirs de l'image (considérés comme non importants pour le diagnostic [99]) sont éliminés. Ainsi, après avoir ignoré les valeurs des coins noirs de l'image, seules 1680 valeurs de hachage sont codées. Enfin, pour choisir une des partitions de la figure 4.4 pour un macro-bloc, nous utilisons l'algorithme suivant :

- 1- Nous regroupons chacun des quatre blocs 4x4 dans un bloc 8x8, comme dans la figure 4.4, et nous calculons sa valeur moyenne.
- 2- La moyenne de chaque bloc 8x8 est comparée à la moyenne de chaque bloc 4x4, en utilisant l'équation (4.1) :

$$Diff = \sum_{i=1}^4 |Mb8 - Mb4(i)| \quad (4.1)$$

Où le $Mb8$ est la valeur moyenne du bloc 8x8 et $Mb4$ est la valeur moyenne du bloc 4x4.

- 3- Nous répétons ce processus pour chaque bloc 8x8.
- 4- Le bloc 8x8 ayant le plus grand $Diff$ n'est alors pas regroupé et reste divisé en blocs 4x4 ;

Pour le codage WZ, d'abord la DCT 4x4 entière est appliquée sur la trame WZ [34]. Ensuite, les coefficients du même ordre k sont regroupés en bandes de fréquences. Pour la DCT 4x4, les bandes sont b_k , $k \in \{0, \dots, 15\}$, et chaque bande est quantifiée et convertie en plans de bits, puis codée par le codeur SW. Ces opérations sont effectuées comme dans le système DVC- HDC.

4.3.2. Décodeur

D'abord, le décodeur reçoit l'information de hachage et l'exploite pour générer la meilleure information adjacente (SI) possible. Ensuite, il effectue le décodage SW pour corriger SI.

4.3.2.1. Création du dictionnaire (CB)

L'anatomie des différentes parties du tractus GI change d'une partie à l'autre, donc nous allons générer un CB pour chaque partie et la rechercher séparément. Il est évidemment possible de générer un CB pour l'ensemble du tractus GI, mais il aura une très grande taille. Dans le présent travail, un CB pour l'intestin grêle est créé hors ligne comme suit :

- Nous recueillons un grand nombre des séquences de la CE.
- Nous classifions tous les trames en fonction de leur emplacement et n'utilisons que ceux

Chapitre 4 : codage vidéo distribué basé sur la quantification vectorielle

du poste GI ciblé (dans notre cas, l'intestin grêle).

- Chaque image, des séquences, est partitionnée en macro-blocs 16x16. Chaque macro-bloc représente un CW possible du CB.
- L'algorithme k-means [114] est utilisé pour créer le CB.

Le CB réalisé est prêt à être utilisé, mais il peut-être trop grand. Nous pouvons utiliser une recherche exhaustive, dans le CB, pour produire le meilleur résultat possible, mais avec une grande complexité de calcul. Pour faciliter la recherche, nous proposons d'organiser le CB comme un arbre (figure 4.5). L'algorithme K-means est de nouveau utilisé pour partitionner le CB en groupes. Tous les groupes obtenus, à partir de cette partition, sont positionnés au niveau 2 (voir figure.4.5). Le CW, qui se trouve dans la position centrale du groupe, est utilisé pour représenter ce groupe et il est positionné au niveau 1. La recherche, dans le CB, commence au niveau 1 pour identifier le meilleur groupe. Ensuite, la recherche est effectuée uniquement dans le groupe sélectionné.

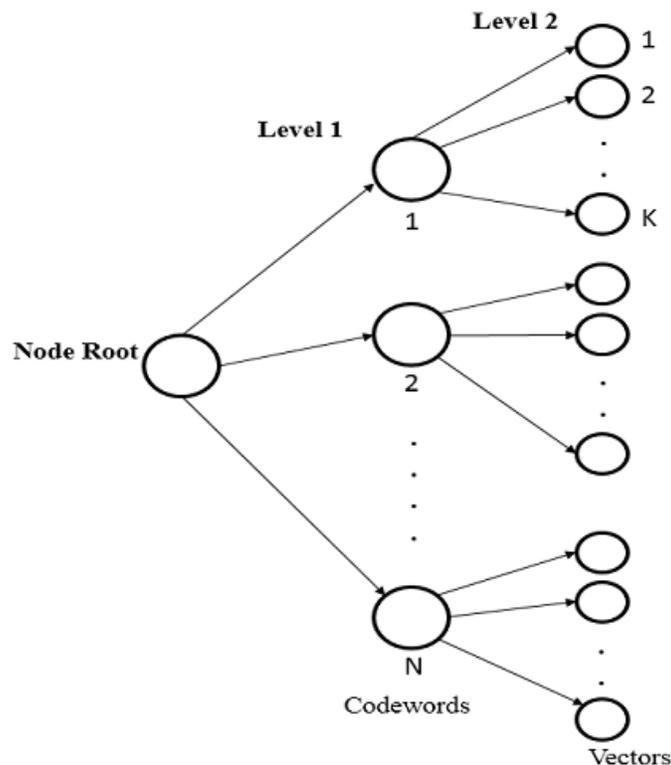


Figure 4.5-L'organisation d'un dictionnaire

4.3.2.2. Génération de l'information adjacente

Le décodeur reçoit les bits de l'information de hachage pour chaque macro-bloc et utilise le décodeur Golomb-Rice et le décodage prédictif pour les décompresser. Pour chaque macro-bloc, une recherche est effectuée dans le CB, en utilisant ce hash pour trouver le meilleur CW. Les CWs sélectionnés sont regroupés pour former l'image décodée, qui sera la SI du DVC.

Avant le décodage SW, des opérations supplémentaires sont utilisées pour améliorer SI. Tout d'abord, l'image décodée de la QV est améliorée par les valeurs de hachage. Parce que, le codeur calcule les valeurs moyennes du bloc 4x4 en tant que hash, les valeurs du hash sont donc meilleures que n'importe quelle valeur moyenne obtenue à partir du CB. Par conséquent, la SI finale est une combinaison de la CW sélectionnée à partir du CB et des valeurs de hachage transmises. Donc, les valeurs moyennes 4x4 du hash sont utilisées directement pour remplacer les valeurs DC de l'image reconstituée de la QV et l'image décodée WZ. Ensuite, le décodeur sait que les pixels en dehors de l'image circulaire sont noirs, donc, il corrige également ces pixels (dans l'image SI) pour réduire le nombre d'erreurs. Enfin, comme les blocs des images SI proviennent d'images différentes, un effet de blocs est possible. La génération de la SI en utilisant les blocs du CB introduit une discontinuité entre les blocs. Une solution simple est d'utiliser un simple filtre passe-bas, mais ce filtre entraîne des pertes de netteté et de qualité. Alors, un filtre adaptatif réducteur des effets de discontinuité, proposé par [115], a été utilisé, pour réduire les erreurs de la SI et améliorer le décodage.

4.3.2.3. Modélisation de la corrélation

Pour améliorer la qualité et le débit au niveau du décodeur, il faut bien modéliser le bruit entre la trame de WZ et la SI. L'efficacité de notre décodage va dépendre de la bonne estimation du paramètre α du modèle Laplacien. Comme vu dans le chapitre précédent qu'il n'est pas possible de calculer le paramètre α du modèle Laplacien à partir de la différence entre les images clés. Pour résoudre ce problème, des solutions sont proposées : soit utiliser une estimation hors-ligne, soit une extrapolation à partir des images WZ déjà décodées. Par conséquent, nous avons testé les deux techniques et nous avons décidé d'utiliser l'estimation hors ligne.

Modélisation de corrélation hors ligne

En fait, l'estimation hors ligne est utilisée uniquement pour les coefficients AC, car pour les coefficients DC, les paramètres peuvent être estimés en ligne. Puisque certaines valeurs DC originales (valeurs moyennes) sont transmises, nous calculons la différence entre elles et les valeurs DC de la SI (sélectionnées à partir de CW). Ces différences sont utilisées pour estimer le modèle d'erreur pour les valeurs DC non transmises. Pour une estimation hors ligne, nous avons retenu les paramètres α , du modèle Laplacien, pour chaque bande DCT, après de nombreux tests. L'estimation hors ligne des paramètres du canal de corrélation a été effectuée comme proposé par [116]. Les étapes suivantes doivent être effectuées pour estimer le paramètre α :

- **Première étape**, on calcule la trame résidu entre la trame WZ et la SI en utilisant :

$$R(x, y) = WZ(x, y) - SI(x, y) \quad (4.2)$$

Où (x, y) est la position du pixel dans la trame WZ et SI.

- **Deuxième étape**, la DCT 4x4 appliquée sur la trame résidu R et nous obtenons les coefficients DCT de la trame résidu,

$$\tau = DCT(R) \quad (4.3)$$

- **Troisième étape**, on calcule la variance σ , mais avant de calculer la variance, on calcule deux termes $E[Z^2]$ et $E[Z]$,

$$\sigma = E[Z^2] - (E[Z])^2 \quad (4.4)$$

Dans ce contexte, le calcul de la variance de la bande DCT contient b coefficients, $E[Z^2]$ et $E[Z]$ correspond à $E[\tau_b^2]$ et $E[\tau_b]$, respectivement, et nous calculons,

$$\begin{cases} E[\tau_b] = \frac{1}{L} \times \sum_{k=1}^L \tau_b(k) \\ E[\tau_b^2] = \frac{1}{L} \times \sum_{k=1}^L (\tau_b(k))^2 \end{cases} \quad (4.5)$$

τ_b Représente les coefficients DCT de b bandes de la DCT de trame résidu et L représente le nombre des coefficients dans la bande DCT.

- **Quatrième étape**, on estime le paramètre α pour chaque bande des AC en utilisant,

$$\alpha_b = \sqrt{\frac{2}{\sigma_b^2}} \quad (4.6)$$

Le b représente le nombre des bandes du DCT.

- **Cinquième étape**, on calcule la moyenne du paramètre α hors-ligne,

$$\alpha_b = \frac{1}{N} \times \sum_{k=1}^N \alpha_{b,k} \quad (4.7)$$

Chapitre 4 : codage vidéo distribué basé sur la quantification vectorielle

Le N représente le nombre des trames hors-ligne utilisé pour générer le paramètre α .

Modélisation de corrélation en ligne

Pendant le décodage, le paramètre α du modèle Laplacien est lu à partir d'un tableau. De plus, la valeur hors ligne est utilisée pour commencer le décodage du premier plan de bits des coefficients AC. Pour les autres plans de bits des coefficients AC, le paramètre α est amélioré, en ligne, après décodage de chaque plan de bits. Cette opération est effectuée en utilisant la méthode de maximum de vraisemblance, proposée par [54]. La méthode d'estimation du paramètre α est donnée par l'algorithme suivant :

Première étape : on calcule S_a et S_b correspondant au nombre des coefficients DCT de l'information adjacente y_i inférieure ou supérieure aux bornes de la quantification u_i et v_i :

$$\mathbf{si}(y_i < u_i) S_a = S_a + 1; \quad (4.8)$$

$$\mathbf{si}(y_i > v_i) S_b = S_b + 1; \quad (4.9)$$

Deuxième étape : on applique l'algorithme suivant :

$$\mathbf{si}(u_i \leq y_i \leq v_i) \quad \forall i \in \{1, 2, \dots, k\}$$

$$\mathbf{alors} \quad \alpha_l^m = \alpha_l^{m-1} \quad (4.10)$$

$$\mathbf{sinon} (y_i \notin [u_i, v_i])$$

$$\mathbf{alors} \quad \alpha_l^m = \frac{1}{\Delta^m} \times \ln \frac{\sum_{i \in S_a} [v_i^{(m)} - y_i] + \sum_{i \in S_b} [y_i - v_i^{(m)}] + |S_b| \Delta^m}{\sum_{i \in S_a} [v_i^{(m)} - y_i] + \sum_{i \in S_b} [y_i - v_i^{(m)}] - |S_a| \Delta^m} \quad (4.11)$$

Où k est le nombre des coefficients AC du DCT, m représente le numéro du plan de bits, α représente le paramètre du modèle Laplacien et Δ^m représente la différence entre les bornes de la quantification: $\Delta^m = v^{(m)} - u^{(m)}$. Le paramètre $\alpha^{(m)}$ de plan de bits est égal au paramètre $\alpha^{(m-1)}$ de plan de bits précédent si $y_i \in [u_i, v_i]$, sinon l'équation (4.11) est appliquée pour calculer $\alpha^{(m)}$.

Troisième étape : si on a dans le plan de bits des coefficients AC de l'information adjacente y_i inférieure et supérieure aux bornes de la quantification on applique cette équation :

$$\alpha_l = \frac{\alpha_l^{m-1} \times [K - (S_a + S_b)] + \alpha_l^m \times (S_a + S_b)}{K} \quad (4.12)$$

Où α_l^{m-1} et α_l^m correspondant au paramètre du modèle Laplacien du plan de bits précédent et le paramètre calculé par équation (4.11).

4.3.2.4. Décodage Slepian-Wolf

Le décodeur WZ utilise la SI (générée à partir de la QV) pour décoder le flux de bits WZ dans le domaine transformé. Par conséquent, nous appliquons la DCT entière sur la trame de SI, puis le décodage commence en demandant les bits du syndrome à partir de l'encodeur. Nous utilisons SI pour décoder les bandes DCT, bande par bande en commençant par le DC. Une fois que tous les plans de bits des bandes DCT quantifiées sont décodés, nous produisons la trame WZ décodée comme dans [117]. Cette opération exploite les valeurs finales des paramètres du modèle Laplacien. Enfin, les coefficients DCT qui sont éliminés par quantification sont remplacés par ceux de l'information adjacente.

4.4. Analyse de la complexité de l'encodeur

Le système proposé a été conçu pour réduire la complexité du codeur, mais avec une complexité accrue du décodeur. Le facteur principal qui a réduit la complexité du codeur est l'utilisation d'un seul type de trame (trame WZ) et l'utilisation d'une information de hachage simple. Il est démontré par [118], [119] et [98] que le codage des trames WZ est plus simple que le codage intra H.264. Cependant, une comparaison directe détaillée de DVC avec JPEG n'est pas disponible, mais les codeurs peuvent être considérés comme ayant une complexité similaire. Le codage WZ et le JPEG sont moins complexes que H.264 [120], car ils effectuent une DCT directe sans prédiction spatiale (comme dans H.264). La DCT de H.264 (utilisée dans notre DVC) est beaucoup plus simple que celle de JPEG, car il s'agit d'une transformation entière qui a été optimisée pour des calculs simples, avec des opérations de décalage. Le codage entropique de H.264 est également complexe, car il utilise la modélisation avec contexte. Cependant, la complexité du codage SW et le codage d'entropie de JPEG sont tous les deux faibles ; avec une performance de compression supérieure pour le codage SW. En outre, des implémentations matérielles de la DCT de 4x4 (utilisés en H.264), des codes Golomb-Rice et LDPC (utilisés dans de nombreux systèmes de communication) sont également disponibles.

Étant donné que la plupart de la complexité de l'encodeur dans les systèmes DVC provient du codage intra des trames clés, le codeur de notre DVC est beaucoup plus simple et plus efficace que les systèmes proposés par [28] et [98], car il utilise uniquement des trames WZ. L'encodage de nos images WZ avec l'information de hachage est plus complexe que celui d'encodage de trames

Chapitre 4 : codage vidéo distribué basé sur la quantification vectorielle

WZ sans l'information de hachage (comme dans [28]), mais la complexité de hachage supplémentaire reste très faible par rapport au codage intra.

Le calcul et le codage de notre information de hachage sont également beaucoup plus simples que celui de [98] qui utilise (en tant que hash) des images sous-échantillonnées et codées intra. En outre, le calcul des valeurs moyennes du hash proposé peuvent même être mis en œuvre en tant que technique de traitement dans le capteur CMOS [121]. En plus, le calcul de notre hash n'a besoin d'aucun traitement supplémentaire, comparé à tout autre type de hash. Ceci est dû au fait que le calcul des valeurs moyennes est identique au calcul des valeurs DC de la DCT, nous pouvons donc obtenir directement les valeurs moyennes à partir de la DCT, sans calcul.

En contrepartie, le décodage WZ est plus complexe que l'encodage en raison de la recherche dans le CB et du décodage LDPC répétitif. Le décodage WZ est également plus complexe que le décodage intra de H.264 [34] ou JPEG. Par conséquent, la complexité du décodeur est augmentée par rapport à celle de [98] aussi, car toutes les images sont WZ. La recherche dans le CB est aussi plus complexe que l'estimation et la compensation de mouvement. L'estimation de mouvement compare chaque bloc d'image seulement avec des blocs de la région de recherche, mais la QV compare le bloc avec tous les candidats CB. En fait, c'est pourquoi nous proposons d'organiser le CB en groupes. Par exemple, pour une estimation de mouvement des macro-blocs dans une zone de recherche de ± 16 pixels, il existe 1039 comparaisons par macro-blocs [122]. Ce nombre est doublé à 2078 pour l'interpolation bidirectionnelle de SI. Ceci est très faible par rapport à la taille de notre CB, mais avec la recherche hiérarchique, le nombre moyen de comparaisons de notre recherche est proche de celui de l'estimation du mouvement. En plus, les comparaisons de hash n'utilisent pas tous les pixels du macro-bloc. Donc, la recherche hiérarchique est un compromis entre la vitesse de décodage et la taille CB.

Du point de vue des ressources, la complexité du décodeur n'est pas le principal problème pour la CE et les autres applications ciblées par codage WZ, car nous avons supposé que le décodeur contient de grandes ressources. Ainsi, le temps de traitement dans le décodeur dépend de l'implémentation matérielle du décodeur. Un décodeur portable peut avoir une batterie et des contraintes de traitement, mais pour les décodeurs non portables, nous pouvons prendre les opérations de décodage en temps réel [60,123]. En fait, la fréquence et la taille d'images sont, actuellement, faibles en raison des limitations de l'encodeur. Si nous augmentons la fréquence d'images de 2 à 10 images/sec, la CE fonctionnera seulement 2 heures au lieu de 10. Par

Chapitre 4 : codage vidéo distribué basé sur la quantification vectorielle

conséquent, le décodeur n'a besoin d'utiliser que les ressources qui lui permettent de fonctionner en temps réel à la vitesse de l'encodeur. Ainsi, pour un temps de décodage donné (latence) et des ressources de calcul du décodeur, nous utilisons le plus grand CB et les techniques de recherche possibles, pour permettre la meilleure performance RD.

Ce n'est que récemment que des solutions DVC avec une capacité de décodage en temps réel ont été proposées, mais elles sont basées sur un traitement parallèle utilisant les GPGPU (*General-purpose computing on graphics processing units*) ; et avec des trames clés et des opérations simplifiées de création de la SI [60,123]. Ces mêmes solutions sont largement suffisantes pour le décodage à 2 images/sec dans la CE. Cependant, la recherche de la QV est plus complexe que l'estimation de mouvement et nous devons donc la réduire pour permettre le décodage en temps réel. Enfin, nous devons mentionner que, dans le DVC, de nombreuses opérations qui affectent la performance RD sont indépendantes de l'encodeur. Ainsi, nous pouvons garder le même codeur simple et faire des changements dans le décodeur pour obtenir une meilleure implémentation matérielle. Encore mieux dans une implémentation logicielle GPGPU, nous pouvons facilement mettre à jour le décodeur avec des versions améliorées des opérations itératives et progressives de création de la SI, et de meilleures opérations de décodage et de modélisation de canal.

D'un autre côté, n'oublions pas que l'encodeur et le décodeur fonctionnent ensemble pendant la même période de temps et si l'un s'arrête l'autre s'arrête aussi. Pour la CE, c'est l'encodeur qui s'arrête, toujours, le premier lorsque la batterie se vide. Cependant, la complexité du décodeur affecte la latence et la taille du tampon. Le décodeur contrôle le débit de codage par le canal de retour. Pour cela, le codeur est forcé de stocker les données WZ, dans un tampon, pendant tout le temps de décodage. L'avantage de notre DVC est qu'il ne nécessite qu'un seul tampon pour stocker la trame actuelle (plus petite) et il n'est pas nécessaire de stocker les trames futures. On peut aussi supposer que le temps de création et de transmission du hash soit négligeable. Ainsi, la latence de notre système est inférieure à celle de [98], car les images sont encodées dans leur ordre d'origine. Ainsi, une fois que les syndromes sont calculés, il n'est pas nécessaire de stocker la trame originale. Donc, nous pouvons stocker la prochaine image et commencer à calculer son hash durant le codage SW de l'image courante.

Le tampon des syndromes a une taille connue et fixe dans le système DVC. En fait parce que le codeur ne connaît pas le nombre de bits requis par le décodeur, il calcule tous les bits syndrome de la trame (dépend de la quantification) et les stocke. Nous pouvons aussi supposer que le

Chapitre 4 : codage vidéo distribué basé sur la quantification vectorielle

décodeur est capable de compléter le décodage pendant le temps d'une image (en fonction de la fréquence d'images). Dans ce cas, nous n'avons pas besoin de modifier (augmenter) la taille du tampon pour stocker les bits pour plus d'une image, mais nous avons besoin des ressources de décodage qui permettent d'effectuer cette opération en temps réel. Si le décodeur termine rapidement son fonctionnement (en demandant un petit nombre de bits), il attendra le début du décodage de la trame suivante (500 ms pour 2 images/sec). Cependant, si le décodeur ne peut compléter le décodage pendant le temps nécessaire, un tampon supplémentaire (de la taille des bits WZ d'une image) est nécessaire pour stocker les syndromes de la nouvelle image. Étant donné que, pratiquement, la plupart des images seront décodées dans le temps, il n'y a pas un besoin réel pour ce tampon supplémentaire.

Il y a une autre solution à ce problème, qui consiste simplement à attendre jusqu'à la fin du décodage d'une image (quel que soit ce temps) et à demander l'image suivante. Cela entraînera une fréquence d'images variable contrôlée par le décodeur. En fait, c'est l'une des raisons pour lesquelles nous utilisons le DVC pour CE. Notre objectif est d'être en mesure d'augmenter la fréquence d'image, et même la taille des trames, lors du passage de la capsule dans une région importante et de la diminuer pour d'autres régions. Ainsi, avec des ressources suffisantes au décodeur, pour fonctionner à des fréquences de trames élevées, le problème deviendra celui de la gestion optimale de l'énergie du codeur.

Le temps pendant lequel les bits de la trame WZ doivent être stockés dans l'encodeur dépend uniquement du temps de génération de la SI (T_{si}) et du décodage SW. Un temps long peut être utilisé par le décodage SW, car cela dépend du nombre de demandes effectuées par le canal de retour (F). Chaque demande comprend le temps pour une transmission supplémentaire de bits de syndrome (T_{tr}) et un autre décodage LDPC (T_{ldpc}). Cette opération dépend d'un facteur supplémentaire de la qualité de l'estimation du débit initial de décodage [124]. Les techniques de traitement en parallèle actuelles, qui sont basées sur les architectures GPGPU, permettent des implémentations en temps réel de cette opération [60] [123] [125] même pour les applications ayant une fréquence d'image beaucoup plus élevée que celle des CE. Dans ce contexte, un avantage du DVC basé sur la QV est qu'il permet la création de la SI et le codage SW en même temps (pipeline); au lieu de créer la SI et ensuite effectuer un décodage SW. Donc, les deux opérations peuvent être réalisées en parallèle. Ainsi, le codeur peut envoyer, en même temps, les bits SW de la trame courante et le hash de la trame suivante (s'il est prêt). Cela permettra au

Chapitre 4 : codage vidéo distribué basé sur la quantification vectorielle

décodeur de générer la SI, de la deuxième image, durant le temps de décodage des bits SW de la première image.

Les opérations complexes de décodage LDPC et la création de la SI peuvent être totalement mises en parallèle et implémentées sur GPGPU ou avec des architectures matérielles, pour permettre un décodage en temps réel. Cependant, le contrôle des débits nécessite un algorithme plus efficace qui réduit F . Dans le cas d'architectures sans canal du retour, nous pouvons facilement atteindre les 30 images par seconde pour l'entrée CIF [126]. Par conséquent, le paramètre principal qui contrôle la latence de notre système (par rapport à JPEG) est F . Plus précisément, pour déterminer la limite supérieure et inférieure de la fréquence d'image de décodage f_r dans le système, il faut calculer le temps de décodage d'une trame WZ qu'on va appeler T_f . On peut dire que T_f est la somme de T_{si} , T_{tr} et T_{ldpc} [98]. Le temps de transmission T_{tr} est très petit et peut-être ignoré (de l'ordre de 15ns [127]). Ainsi, le délai de décodage est principalement dû à T_{si} et T_{ldpc} [128] et dans notre système, T_{si} est le temps de la recherche dans le CB. Dans ce cas, la fréquence d'images est calculée comme suit :

$$f_r = \left[\frac{1}{T_{si} + F \times T_{ldpc}} \right] \quad (4.13)$$

Dans l'analyse qui suit, nous utilisons certaines valeurs pratiques de ces temps de traitement que nous avons trouvés dans la littérature. Dans [129], les auteurs ont implémenté un codec QV en temps réel, pour les films. Le processeur QV recherche un CB qui contient 2048 vecteurs en seulement 780 ns et peut ainsi coder 30 images/sec. Cette valeur peut être similaire à notre mode de recherche d'arborescence avec 3500 vecteurs. Alternativement, dans [98], les auteurs ont calculé le temps d'estimation de mouvement pour une image 256x256, en fonction du travail de [122], à 10,63 ms. Dans notre analyse, nous allons considérer le pire des cas et supposer un T_{si} similaire à leur valeur.

Pour T_{ldpc} , les auteurs de [130] ont calculé le temps de décodage LDPC avec une longueur de mot de code de 64800 bits (norme DVB-S2) avec 50 itérations et ont obtenu un temps de 270 μ s. Dans notre système, nous utilisons une longueur de code de seulement 3762 bits et 80 itérations et nous considérons encore le pire scénario d'un T_{ldpc} égale à 270 μ s également. Ainsi, notre équation (2) deviendra :

$$f = \left[\frac{1}{(10.63 + F \times 0.270) \times 10^{-3}} \right] \quad (4.14)$$

Pour $F = 1$, ce qui signifie que tous les syndromes sont transmis une fois (équivalant à un codage sans canal de retour), la fréquence d'image f_r est de 91 images/sec. Pour le pire des cas, nous considérons que le décodeur va demander tous les paquets de syndromes, en commençant par transmettre un seul paquet. Il existe 64 plans de bits (pour la plus grande matrice de quantification) avec LDPCA de 66 paquets, ce qui entraînera $F = 4224$. Dans ce cas, $f_r = 0.8687$ images/sec. Cependant, les auteurs de [98] ont calculé le nombre des demandes du canal, d'un décodeur WZ, pour la capsule endoscopique pour différentes matrices de quantification et le nombre maximum est inférieur à 100. Pour cette valeur de F , la fréquence d'images f_r est 26 images/sec. Bien que, dans notre cas, il n'y ait pas d'images clés comme dans [98], nous pouvons utiliser une valeur similaire de F . Cette valeur dépend uniquement de la qualité de la SI, meilleure pour DVC basé sur VQ, et de l'algorithme de contrôle de débit. Néanmoins, si nous considérons le pire des cas de $F = 1000$, la fréquence d'images f_r sera de 3,56 images/sec. Ces valeurs peuvent être augmentées si LDPCA et la création de la SI sont exécutées en parallèle.

4.5. Résultats expérimentaux

Dans cette section, nous évaluerons le système DVC proposé (que nous appelons DVC-VQ) en présentant la qualité de son SI, sa performance RD et ses possibilités de gestion de l'énergie. La performance est mesurée en termes de valeurs moyennes du débit et de PSNR des séquences codées. Le débit total est la somme de bits de l'information du hachage et de bits WZ. Les résultats sont comparés à ceux de MJPEG [106] et intra H.264 [64]. Pour montrer les améliorations obtenues (par l'utilisation de QV), nous comparons également les résultats du DVC-VQ avec ceux du système DVC proposé dans le chapitre précédent (DVC-DC). Dans nos tests, nous avons utilisé quatre vidéos de la CE de l'intestin grêle, chacune contenant environ 50 images. Ces séquences sont: "test vidéo1" qui contient des trames normales et claires de l'intestin grêle, "test vidéo2" qui contient une séquence complexe, prise à partir d'une CE qui a des mouvements irréguliers, "test vidéo3" qui contient des images avec des bulles d'air et "test vidéo4" qui représente des images moins claires de l'intestin grêle. Le débit d'acquisition de la CE est de 2

Chapitre 4 : codage vidéo distribué basé sur la quantification vectorielle

images par seconde avec une résolution d'image de 256x256 pixels. Pour mesurer l'efficacité de la technique proposée, nous avons effectué les tests suivants :

4.5.1. La qualité de l'information adjacente

Dans cette analyse, nous montrerons l'effet de l'utilisation de la QV sur la qualité de la SI générée. Nous rappelons que la qualité de la SI est un élément clé de la performance du système DVC-VQ. En particulier, nous verrons les performances de l'algorithme de recherche hiérarchique par rapport à l'algorithme de recherche complète.

Tout d'abord, pour créer le CB, nous avons utilisé un groupe de 136 séquences vidéo qui n'incluent pas les séquences de tests. La partition de ces séquences donne 3.389.440 blocs (CW possibles). Après avoir appliqué l'algorithme k-means, nous avons obtenu un CB avec une taille de 96 000 CW (niveau 2 de la figure 4.5). C'est notre plus grand CB et deux autres CB avec des tailles plus petites de 48000 et 24000 CW ont été créés également. Pour accélérer la recherche, nous avons organisé le plus grand CB en 3400 groupes (niveau 1 de la figure 4.5), en utilisant à nouveau l'algorithme K-means. Chaque groupe, du niveau 1, regroupe un nombre différent de CW, variant de 33 à 1000.

La figure 4.6 montre la qualité des trames de la SI mesurée par le PSNR de chaque image. Dans la figure 4.6, nous comparons la qualité de la SI générée par le CB de 96000 CWs avec recherche exhaustive et recherche hiérarchique. La figure 4.6 montre aussi la qualité de la SI pour deux CB de tailles 48000 et 24000 CWs en utilisant la recherche exhaustive. On remarque que la qualité de la SI est toujours améliorée en augmentant la taille du CB, même si elle ne double pas en doublant la taille du CB. Étant donné que le hash est calculé indépendamment de la taille du CB, nous pouvons utiliser le CB avec la plus grande taille en fonction de la complexité du décodeur et de sa latence. D'autre part, comme on pouvait s'y attendre, nous pouvons constater que la qualité de la SI qui a été générée par la recherche exhaustive est toujours meilleure que celle de la recherche hiérarchique.

La qualité de la recherche hiérarchique est semblable à celle de l'utilisation des plus petits CB, mais la recherche hiérarchique est beaucoup plus rapide. Alors que l'utilisation des petits CB est seulement 2 ou 4 fois plus rapide, la recherche hiérarchique était 20 fois plus rapide que la recherche exhaustive.

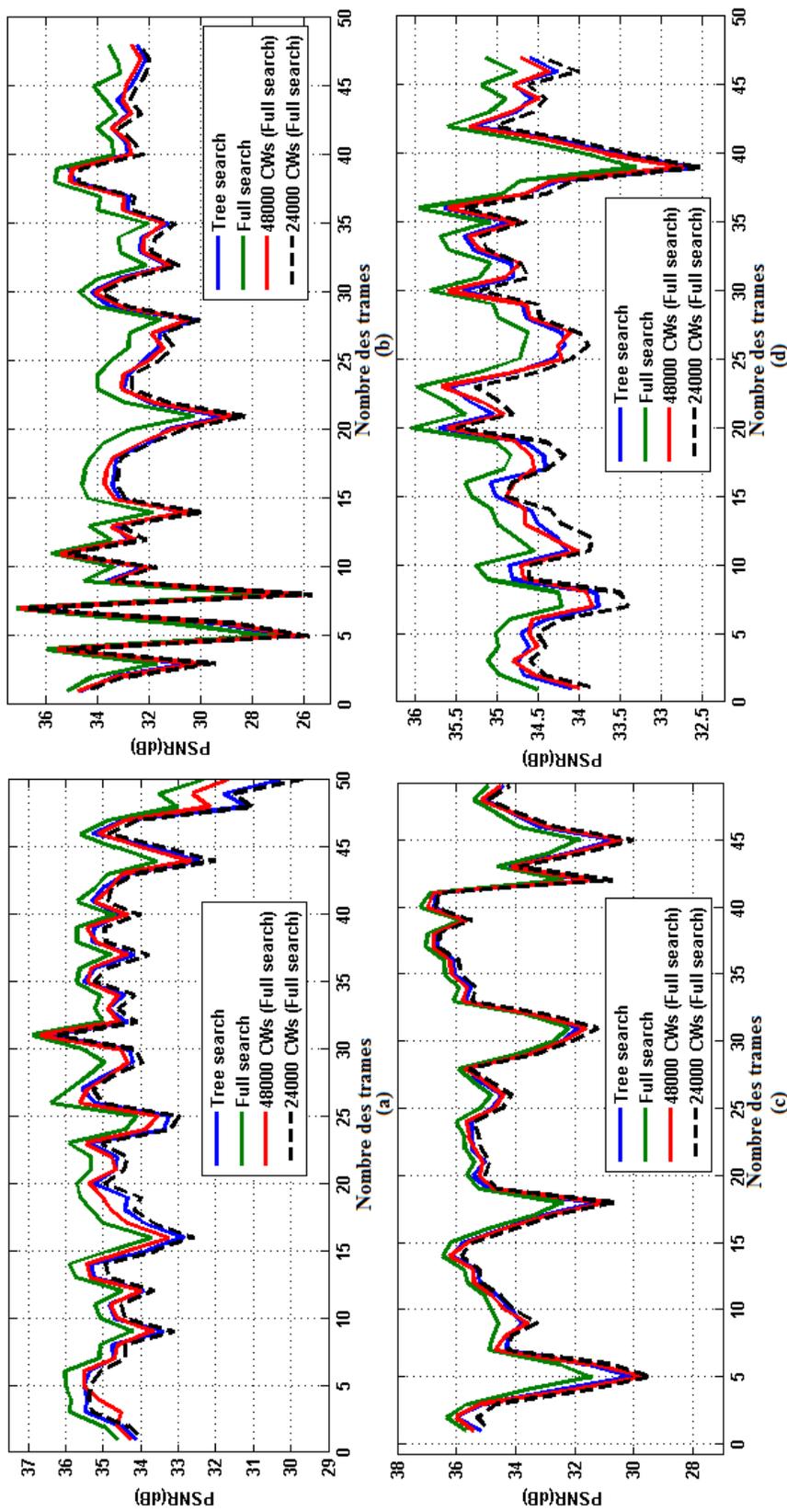


Figure 4.6- La qualité de l'information adjacente avec une recherche complète et une recherche arbre : (a) "test vidéo 1", (b) "test vidéo 2", (c) "test vidéo 3" et (d) "test vidéo 4".

Chapitre 4 : codage vidéo distribué basé sur la quantification vectorielle

Le tableau 4.1 présente certaines statistiques des résultats des images de la SI obtenue. Le PSNR moyen est supérieur à 32 dB pour toutes les séquences de test, ce qui est déjà une bonne qualité. Les trames de la SI avec une mauvaise qualité sont rares et ils sont proches de la moyenne. Ce problème peut-être (la plupart du temps) causé par quelques blocs pour lesquels il n'y a pas de bonne correspondance dans le CB. Par exemple, la qualité moyenne de la SI, dans le test vidéo2, est la plus faible parmi les vidéos de test. Dans cette séquence, il existe des trames très complexes et le décodeur ne peut pas trouver de bons blocs de la SI correspondants pour eux, dans le CB. Par conséquent, nous pouvons voir une grande variation de sa qualité.

Avec DVC-VQ, la qualité de la SI de chaque image WZ, dans les séquences de la CE, peut-être indépendante de la qualité des images précédemment décodées et de leur SI. Généralement, les images proches ont une qualité de SI similaire, mais parfois, de grands changements sont observés. Cela peut-être une indication d'un changement de scène en raison d'un mouvement brusque de la CE ou de l'apparition d'un nouveau contenu dans la scène. Néanmoins, le DVC-VQ n'est pas largement affecté par ces phénomènes et la qualité de chaque image dépend principalement de son propre contenu et de celui du CB. Par conséquent, cela signifie que DVC-VQ est plus robuste au

Tableau 4.1- Les propriétés de la qualité de l'information adjacente de la figure 4.6.

Test vidéo	PSNR moyen(dB)	Valeur maximum (dB)	Variance
Test vidéo 1 (recherche d'arbre)	34.42	36.68	1.36
Test vidéo 1 (recherche totale)	35.10	36.88	0.72
Test vidéo 2 (recherche d'arbre)	32.42	36.97	4.01
Test vidéo 2 (recherche totale)	33.37	37.14	3.36
Test vidéo 3 (recherche d'arbre)	34.44	36.95	2.80
Test vidéo 3 (recherche totale)	34.90	37.20	2.14
Test vidéo 4 (recherche d'arbre)	34.64	35.71	0.33
Test vidéo 4 (recherche totale)	35.02	36.05	0.26

Chapitre 4 : codage vidéo distribué basé sur la quantification vectorielle

mouvement de la CE et aux problèmes de communication, par rapport au DVC basé sur la prédiction. En fait, lorsqu'une image donnée est perdue, en raison d'un problème de communication ou si elle contient des erreurs, l'image suivante est décodée indépendamment de ce qui est arrivé à cette image. Cependant, ignorer la corrélation entre les images proches n'est pas toujours bon pour DVC-VQ, car l'exploitation de ces images peut améliorer ses performances RD. Par conséquent, nous avons décidé d'exploiter l'image préalablement décodée, correctement, pour améliorer la SI. Pour un bloc donné, le bloc de la même position, dans l'image précédente, est également testé avec le hash en plus de ceux du CB. Cela améliorera les résultats dans le cas des séquences à très faible mouvement. En outre, comme le système utilise un décodage progressif, le décodeur peut détecter toute corrélation temporelle entre les images et peut l'exploiter pour améliorer la SI ou pour en arrêter le décodage, selon l'importance de l'image.

En plus de la robustesse, nous nous attendions à ce que DVC-VQ ait une meilleure qualité de la SI que la prédiction interpolée. Pour tester cette affirmation, nous avons comparé le PSNR moyen de la SI de DVC-VQ au PSNR de la SI créé par interpolation. Le tableau 4.2 montre la comparaison entre la qualité de la SI de DVC-VQ (recherche complète et hiérarchique) et l'algorithme d'interpolation à compensation de mouvement proposé par [28]. Les résultats montrent que la qualité de la SI, générée par QV, est largement supérieure à celle de l'interpolation. Les résultats de l'interpolation sont obtenus grâce à notre propre implémentation de l'algorithme de [28] qui nous permet de coder des images de n'importe quelle taille. Notre algorithme produit des résultats

Tableau 4.2 -La qualité moyenne de l'information adjacente

	Information adjacente avec recherche complète	Information adjacente avec recherche orienté	L'information adjacente (DISCOVER [24], GOP 2)
PSNR moyen (dB) (Test vidéo 1)	35.10	34.42	27.98
PSNR moyen (dB) (Test vidéo 2)	33.37	32.42	26.45
PSNR moyen (dB) (Test vidéo 3)	34.90	34.44	24.33
PSNR moyen (dB) (Test vidéo 4)	35.02	34.64	26.31

Chapitre 4 : codage vidéo distribué basé sur la quantification vectorielle

similaires à ceux de [28] pour les séquences de test vidéo populaires de haut débit (15 et 30 images/sec) où le mouvement entre les images est faible. Cependant, dans les séquences CE, où le débit est de 2 images/sec, le mouvement entre certaines images est grand et la méthode d'estimation de mouvement, utilisée pour créer SI, n'est pas efficace.

La figure 4.7 montre la qualité visuelle de la SI générée, pour certaines images sélectionnées. La figure 4.7 montre également le PSNR pour les images de test. Sur la figure 4.7 (a et b), nous pouvons voir que les images obtenues ont une très bonne qualité visuelle. Cependant, ce n'est pas le cas sur la figure 4.7 (c et d), où les discontinuités entre les blocs montrent que les constructions de certains blocs de la SI étaient mauvaises. Cela s'explique par le fait que dans le premier cas, les trames appartiennent à des séquences simples et claires, tandis que dans le second cas, les trames contiennent des blocs peu clairs avec des bulles. Un bloc contenant une texture normale du petit

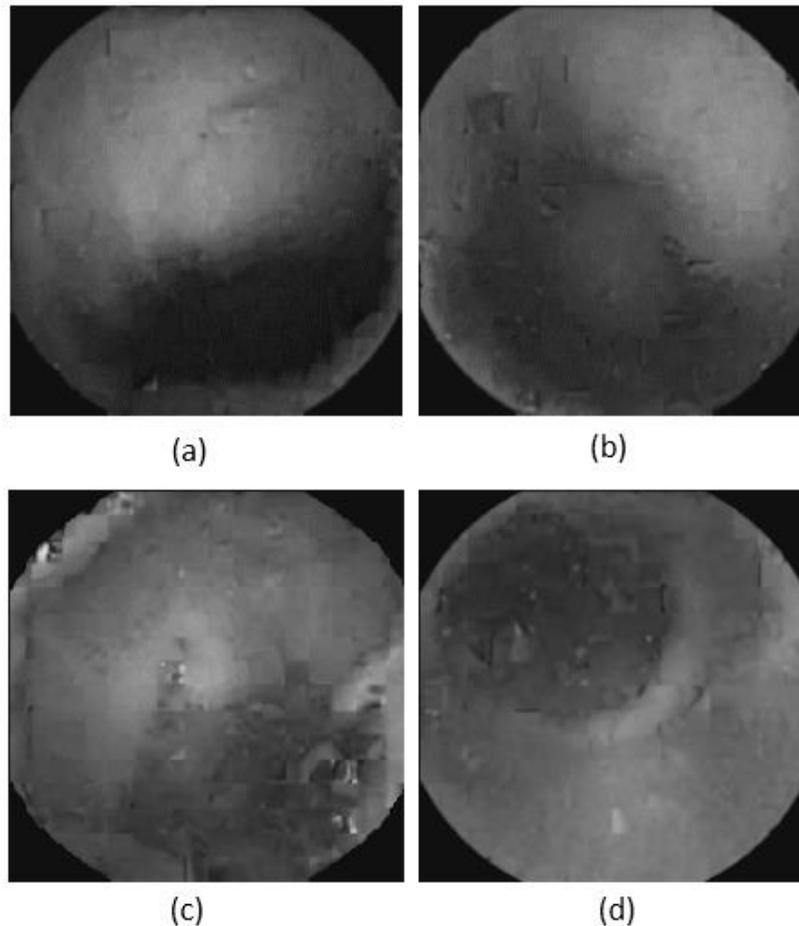


Figure 4.7- La qualité visuelle de l'information adjacente. Les PSNR des trames SI (a), (b), (c) et (d) sont : 34,54 dB, 33,68 dB, 30,11 dB et 32,6 dB.

Chapitre 4 : codage vidéo distribué basé sur la quantification vectorielle

intestin a une forte probabilité d'être répété dans toutes les séquences, donc il sera présent dans le CB. Cependant, un bloc qui contient toute information due à d'autres objets qui ne font pas partie de l'anatomie humaine normale, comme les bulles d'air et les tumeurs, n'aura pas de bons résultats dans le CB. Ce phénomène n'est pas un problème pour le système DVC-VQ, car il entraînera automatiquement une division optimale du débit utilisé entre les images et donc de l'énergie. En d'autres termes, les images normales utiliseront un petit débit et les images anormales consommeront plus de débit. Étant donné que l'objectif initial de l'utilisation de la CE est de rechercher des problèmes médicaux, les images normales sont moins importantes et, par conséquent, consommeront moins d'énergie.

Enfin, l'observation la plus importante est que les images de la SI du DVC-VQ sont déjà de bonne qualité et peuvent parfois être utilisées directement sans amélioration avec des bits SW. C'est le cas lorsque la qualité des images CE reconstituées est supérieure à 35 dB [131]. Cela offre la possibilité de conserver l'énergie et de l'utiliser pour décoder plus d'images.

4.5.2. La performance de codage du système

Pour comparer les résultats RD obtenus, avec MJPEG, nous avons dessiné quatre points RD correspondants aux matrices de quantification 1, 2, 7 et 8 de [27]. Dans ce paramètre expérimental, seul la composante luminance des trames a été codée. Les résultats sont présentés à la figure 4.8. Ces résultats montrent que le codec DVC-VQ proposé dépasse MJPEG [106,135] en permanence. Cependant, il est encore moins efficace que H.264 [34], qui est également plus complexe à utiliser avec les CE. Les auteurs de [132] ont proposé d'utiliser une version modérée peu complexe de H.264 pour le codage dans les CE, mais sa performance était, de manière significative, inférieure. En outre, le codeur H.264 ne peut pas être contrôlé par le décodeur, pour une gestion efficace de l'énergie. Toutefois, ce résultat indique également que les images codées par DVC contiennent toujours une redondance qui n'a pas été réduite et qui doit être supprimée. Dans DVC, cette opération est possible en ajoutant de nombreuses améliorations possibles au niveau du décodeur. Comme prévu, la performance RD du système DVC-VQ avec une recherche exhaustive est meilleure que celle de DVC-VQ avec la recherche hiérarchique. Car, la recherche exhaustive donne la meilleure SI, il est normal qu'il ait la meilleure performance RD. La différence entre la recherche exhaustive et la recherche hiérarchique est plus grande que 0.42dB, mais pour les débits élevés, cette même différence diminue à 0.21dB. Pour les débits faibles, le système DVC-VQ avec

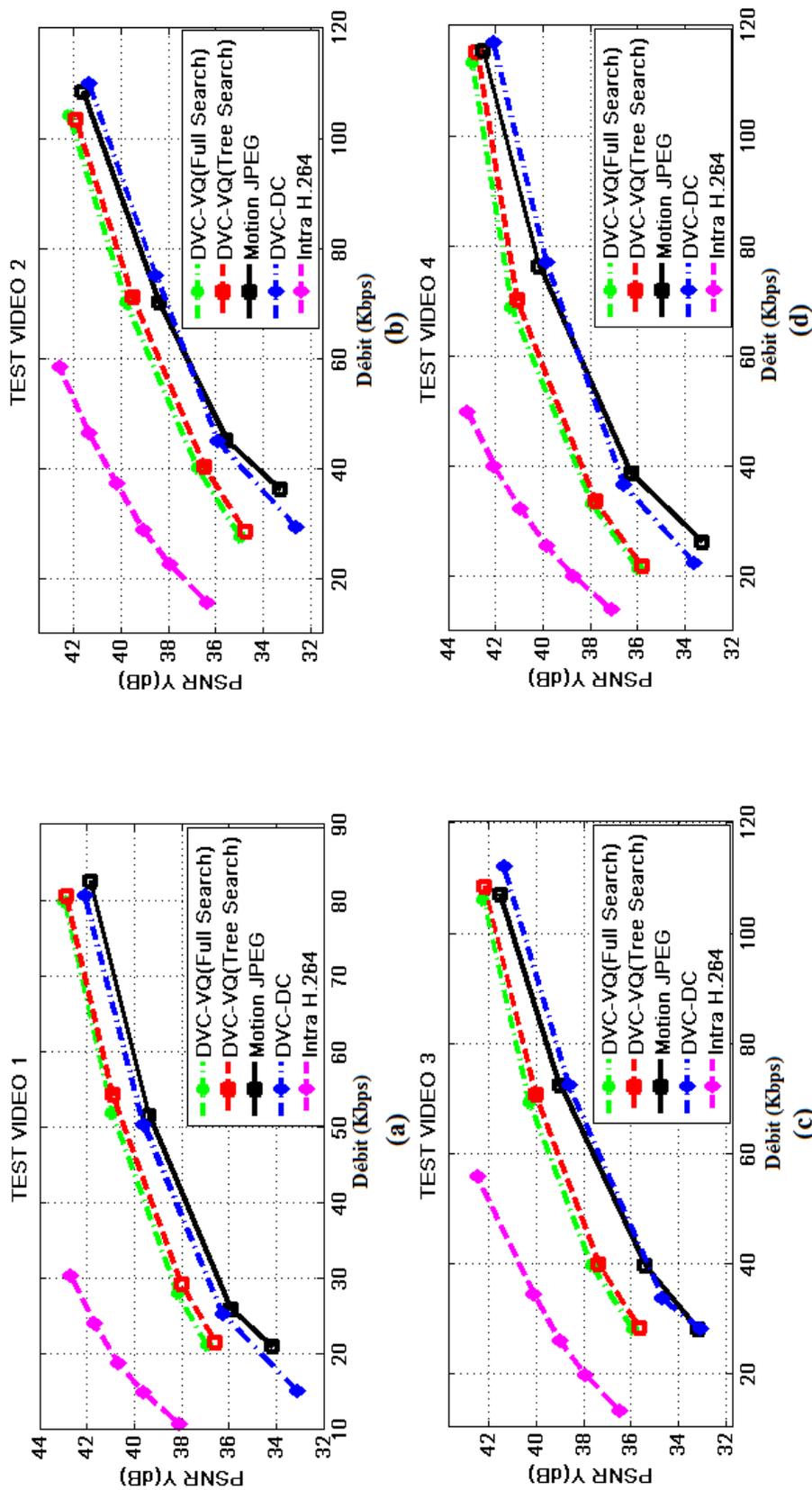


Figure 4.8- Les courbes de la performance RD des séquences (a) “Test vidéo 1”, (b) “Test vidéo 2”, (c) “Test vidéo 3” et (d) “Test vidéo 4” (256x256 pixels avec 2 images/Sec).

Chapitre 4 : codage vidéo distribué basé sur la quantification vectorielle

recherche hiérarchique dépasse largement le MJPEG et le système DVC-HDC de [87]. La différence maximale peut être de 3,4 dB comme le montre la figure 4.8 (d). Le débit moyen de Bjontegaard [108] du codec DVC-VQ est meilleur que celui de MJPEG avec 33,01% sur la Fig. 8 (a), 26,83% sur la Fig. 8 (b), 32,53% sur la Fig. 8 (c) et 51,49% dans la figure 4.8 (d). Le fait que DVC-VQ ne soit que légèrement mieux que MJPEG dans des débits élevés signifie que le système a consommé plus de bits pour le codage SW dans la correction des composants à haute fréquence. La SI est la même pour tous les points RD. Elle est produite par la même information de hachage qui ne coûte qu'environ 7500 bits, ce qui représente 60% de bits pour un faible débit, mais seulement 15% pour un débit élevé. Cela signifie que la SI a plus d'erreurs dans les bandes des hautes fréquences. En fait, la SI est la qualité de base pour tous les points RD et chaque point RD est une amélioration du précédent, puisque chaque point ajoute de nouvelles bandes corrigées ou des plans de bits. Par conséquent, cette évolutivité et la très haute qualité, même à de faibles débits,

Tableau 4.3 -La comparaison entre les modèles d'estimation d'erreur utilisant le mode on ligne et hors-ligne.

Vidéo test 1	α hors ligne ([93])	α on ligne (Extrapolation)	α hors ligne + on ligne	Vidéo test 2	α hors ligne ([93])	α on ligne (Extrapolation)	α hors ligne + on ligne
RD 1 débit (kbps)	10.64	10.59	10.60	RD 1 débit (kbps)	13.88	13.89	13.84
RD 2 débit (kbps)	14.65	14.63	14.65	RD 2 débit (kbps)	20.05	20.11	20.05
RD 3 débit (kbps)	27.26	27.28	27.24	RD 3 débit (kbps)	35.16	35.24	35.20
RD 4 débit (kbps)	40.28	40.31	40.27	RD 4 débit (kbps)	52.13	52.22	52.16
Vidéo test 3	α hors ligne ([93])	α on ligne (Extrapolation)	α hors ligne + on ligne	Vidéo test 4	α hors ligne ([93])	α on ligne (Extrapolation)	α hors ligne + on ligne
RD 1 débit (kbps)	14.06	14.09	14.07	RD 1 débit (kbps)	10.78	10.76	10.78
RD 2 débit (kbps)	19.84	19.91	19.85	RD 2 débit (kbps)	16.66	16.57	16.66
RD 3 débit (kbps)	34.69	34.71	34.70	RD 3 débit (kbps)	34.41	34.52	34.41
RD 4 débit (kbps)	53.12	53.27	53.15	RD 4 débit (kbps)	56.57	56.69	56.57

Chapitre 4 : codage vidéo distribué basé sur la quantification vectorielle

rendent DVC-VQ le meilleur à utiliser pour une détection précoce d'images redondantes et non utiles.

Les résultats des débits ci-dessus sont obtenus en utilisant un modèle hors ligne de statistiques d'erreurs. Avec ce modèle hors ligne, nous réduisons le temps de décodage (latence) sans affecter les performances de RD. Pour vérifier cela, nous avons comparé les résultats de codage de notre modèle hors ligne avec ceux d'une estimation en ligne, des paramètres du modèle, en utilisant une prédiction par extrapolation. Le tableau 4.3 présente les résultats de cette comparaison.

Pour la plupart des points RD, les paramètres hors ligne sont plus efficaces. En fait, tous les résultats sont très proches, il est donc préférable d'utiliser le modèle hors ligne pour réduire la latence. Si la latence n'est pas un problème, la meilleure solution est de combiner l'estimation hors ligne et en ligne des modèles d'erreur. Lorsque les images consécutives sont similaires, l'extrapolation permet une meilleure estimation, mais lorsque les images sont différentes, l'estimation hors ligne assurera un bon résultat.

Tests de contrôle d'énergie

La figure 4.9 montre la qualité visuelle des trames de la SI (a) et l'amélioration progressive des images décodées de DVC-VQ, avec quatre débits différents : (b), (c), (d) et (e). Visuellement, nous pouvons voir que certaines parties sont déjà claires dans la SI et que d'autres sont de bonne qualité juste après le décodage des premiers coefficients DCT. Cependant, certains détails fins de la texture (petites veines sanguines, par exemple) ne sont observés qu'après le décodage de tous les coefficients DCT et la qualité suffisante dépend du contenu de l'image. Il est nécessaire d'avoir un système qui détermine automatiquement le meilleur point pour arrêter l'amélioration. Par exemple, dans la figure 4.9, si l'on considère un PSNR de 40 dB comme un seuil, le décodage de la première séquence peut s'arrêter à la trame (d) sans compléter les plans ou les bandes restantes, tandis que pour la quatrième séquence on doit continuer à (e). Ainsi, dans le premier cas, un débit de 12,52 kbps est enregistré. Le problème est que cette opération n'est possible que si nous pouvons mesurer la qualité du décodeur (sans référence). Ceci est également possible dans une application en temps réel, où les médecins déterminent la qualité et l'utilité de l'image. Cette possibilité, cependant, aidera à résoudre le problème du budget énergétique limité pour la CE. Considérons le problème suivant : La CE prend 24 heures pour parcourir l'intégralité du circuit gastro-intestinal et la batterie est suffisante pour prendre seulement 8 heures d'images de mêmes

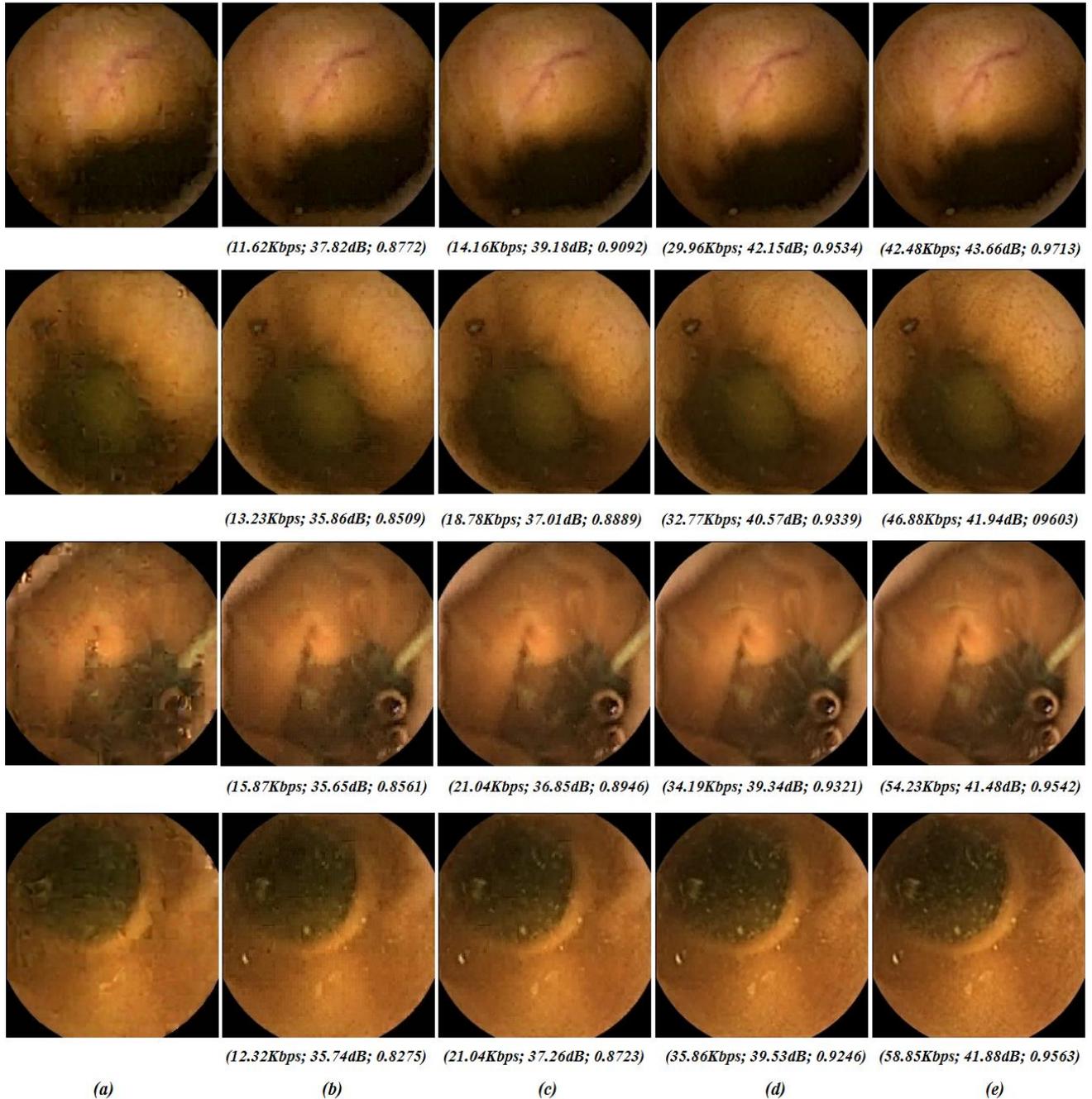


Figure 4.9 - La performance visuelle avec un PSNR de luminance et l'indice SSIM pour "Test Video1", "Test Video2", "Test Video3" et "Test Video4" : (a) trame de l'information adjacente, (b), (c), (d) et (e) trames décodées par DVC-VQ avec des débits différents.

Chapitre 4 : codage vidéo distribué basé sur la quantification vectorielle

qualités et de même fréquence d'image. Donc, comment allouer l'énergie limitée de la batterie afin qu'elle dure 24 heures en codant la meilleure combinaison d'images de bonne qualité (haute énergie) et d'images de faible qualité (faible énergie). Dans DVC, cette opération peut être effectuée par le décodeur (même complexe) qui peut décider de décoder des images de mauvaise qualité (un petit nombre de bits SW) ou sauter l'image complètement en fonction uniquement de la SI (hash uniquement dans le cas de DVC_VQ). Pour tester cette possibilité, les deux expériences suivantes ont été effectuées.

Dans la première expérience, nous avons considéré le cas lorsque la capsule se déplace lentement. Dans notre test, le décodeur calcule le SAD entre les valeurs de hash de la SI courante et celle de l'image précédente. Si le SAD est inférieur à un seuil $Th = 10000$, il arrête de décoder l'image actuelle et la remplace par la précédente. Dans cette démonstration simple, nous avons fixé la valeur de Th après de nombreuses comparaisons de la qualité visuelle des images décodées. Dans un test réel, une séquence de 50 images a été codée et les résultats indiquent que 8 images ont été sautées. Cela signifie que 16% des trames ont été sautées, mais chaque image à un coût différent ; les mesures indiquent que 21% des débits ont été économisés. Si l'on considère que la consommation d'énergie est proportionnelle au débit, cela signifie aussi 21% d'économie d'énergie ou encore une augmentation de la durée de vie de la CE de 21%.

Dans la deuxième expérience, nous avons considéré seulement le codage d'une région d'intérêt. Car, les médecins ne ciblent généralement qu'une petite région du tractus gastro-intestinal. En outre, une capsule qui prend deux images par seconde pendant 8 heures produira 50000 images, mais 10000 ne peuvent comporter aucune information clinique [133]. La sélection d'une région donnée peut-être contrôlée par la fonction de détection de position de certains CE [134] ou par le temps passé par la CE dans le tractus GI.

Dans cette expérience, nous avons utilisé une séquence d'images CE qui contient un saignement intestinal (région d'intérêt). Nous avons contrôlé facilement le débit utilisé avec la quantification. La figure 4.10 montre le PSNR de la SI et les trames décodées de cette séquence codée par DVC-VQ. La qualité de la SI moyenne est de 31,20 dB avec un maximum qui a atteint 34,16 dB et la qualité minimale est de 28,81 dB. Après avoir corrigé la SI, la première partie des images est codée avec une qualité moyenne (en utilisant la deuxième matrice de quantification de [28]) de 37,85 dB avec un maximum de 39,42dB. Pour la deuxième partie des images (région d'intérêt avec les

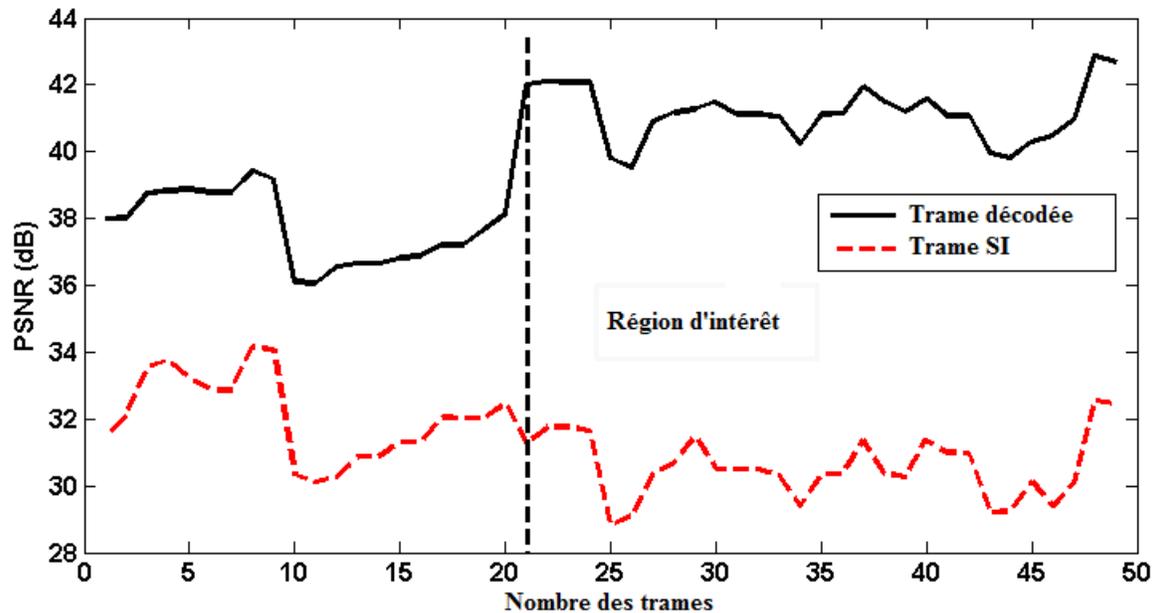


Figure 4.10 - Exemple de codage d'une région d'intérêt avec une haute qualité et une région non informative avec une basse qualité.

trames saignantes), le système utilise la dernière matrice de quantification de [28]. Dans ce cas, la qualité moyenne est de 41,15 dB et le maximum est de 42,88 dB. Le débit moyen dans la région d'intérêt est de 56,2 kbps alors qu'il ne représente que 15,12 kbps pour l'autre région. Pour la séquence totale, le débit moyen est seulement de 38,12Kbps. Par conséquent, nous avons économisé environ 30% du débit par rapport à celui du codage de l'ensemble de la séquence avec une qualité élevée.

4.6. Conclusion

Dans ce chapitre, nous avons introduit un nouveau système de codage vidéo distribué il contient un codeur très simple, qui donne une bonne compression et une bonne qualité d'images décodées. Il combine certaines parties de l'état de l'art du DVC et la quantification vectorielle. Le décodeur utilise l'information de hachage pour rechercher dans un CB pour générer la SI. La qualité de la SI est très bonne et peut atteindre 37 dB, dans certains cas. Les résultats expérimentaux montrent que le codec proposé fournit une compression élevée et qu'il surpasse le MJPEG. Ce système est intrinsèquement évolutif et totalement contrôlé par le décodeur, ce qui permet de contrôler efficacement le débit utilisé.

Conclusion générale

L'objectif de cette thèse est de proposer un nouveau système de codage vidéo distribué pour la capsule endoscopique. Le codage vidéo distribué est un nouveau paradigme de la compression et est un candidat pour les applications sans fil. Son principe est basé sur deux théorèmes Slepian-Wolf et Wyner-Ziv proposés dans les années 70. L'architecture du codec proposé pour la capsule endoscopique doit être plus simple que possible pour réduire la consommation d'énergie dans la capsule, et plus performante par rapport aux outils de la compression traditionnelle qui existe dans les capsules, comme Motion JPEG.

Dans cette thèse de doctorat, on a présenté trois contributions originales du codage Wyner-Ziv vidéo développées dans le chapitre 3 et le chapitre 4. Les systèmes proposés contiennent un encodeur simple, qui utilise un seul type de trame, et utilisent une information de hachage pour la génération d'une information adjacente (SI). Le codage de l'information de hachage utilisé dans nos systèmes n'augmente pas la complexité de l'encodeur, parce que cette information est donnée soit par des coefficients DC déjà calculés ou des valeurs moyennes d'un bloc. Ceci donne un encodeur simple et le plus efficace pour la capsule endoscopique.

Dans le chapitre 3, deux contributions ont été présentées pour la capsule endoscopique en utilisant une base de données pour la génération de l'information adjacente. Ce chapitre explique l'idée de la similarité du corps humain utilisé dans les systèmes proposés. La première contribution a utilisé les coefficients DC déjà calculés par la DCT de trame WZ pour les utiliser comme une information de hachage. Le premier système cherche dans la base de données en

utilisant les coefficients DC. L'image sélectionnée à partir de la base de données est utilisée directement comme une information adjacente pour décoder l'image source. Les résultats obtenus par la première contribution dépassent le Motion JPEG seulement lorsque la compression est haute. Le deuxième système proposé utilise une image sous-échantillonnée comme une information de hachage. Cette information est codée par un codec intra simple comme le JPEG. Le système utilise l'image sous-échantillonnée pour chercher dans la base de données et de sélectionner les images les plus corrélées de l'image source, et les utiliser pour générer l'information adjacente. Ensuite, l'information de hachage et les images sélectionnées à partir de la base de données sont utilisées pour la génération de l'information adjacente. Les résultats obtenus sont meilleurs par rapport à la première contribution parce que la qualité de l'information adjacente générée par la deuxième contribution est meilleure que celle de la première contribution. Aussi, la deuxième contribution est meilleure que le Motion JPEG lorsque la compression est haute, et moins meilleure lorsque la compression est faible. Pour surpasser le Motion JPEG, il faut améliorer la qualité de l'information adjacente au maximum et améliorer l'efficacité de décodage.

Le chapitre 4 a présenté un nouveau système de DVC pour les images de la capsule endoscopique. Le système DVC proposé est basé sur l'idée de la similarité du corps humain. Ce système DVC utilise la technique de la quantification vectorielle pour la génération de l'information adjacente. Le nouveau DVC utilise une nouvelle architecture de la quantification vectorielle qui contient un encodeur simple. Dans notre travail, la quantification vectorielle pour générer l'information adjacente, est utilisée pour décoder la trame WZ. Pour améliorer le décodage, on implémente des algorithmes de la distribution Laplacien qui génèrent des paramètres du système en ligne après chaque plan de bits décodé. Les résultats obtenus montrent la très bonne qualité de l'information adjacente et de la performance RD. En comparaison avec le Motion JPEG, notre système DVC est plus performant que ce soit pour une haute ou faible compression. Le système DVC proposé est simple et efficace pour les images de la capsule endoscopique ou les images médicales en générale. Avec cette contribution, l'objectif principal de la thèse a été atteint en proposant un nouveau système de DVC simple, efficace et applicable pour la capsule endoscopique.

Le codage source distribué est devenu un domaine intéressant dans les codages des applications sans fil. Plusieurs aspects de la recherche sont lancés dans ce domaine comme les réseaux capteurs sans fil, la vidéo de surveillance et le codage de contenu d'écran.

Dans les travaux futurs, on essayera de proposer de nouveaux systèmes de codage pour certaines applications sans fil en utilisant le codage source distribué. Ensuite, on améliorera les systèmes par la proposition des nouvelles techniques de génération de l'information adjacente ou par l'amélioration du décodage. Un autre projet est proposé pour les images médicales, qui consistent à utiliser le codage vidéo distribué en mode sans perte.

Bibliographie

- [1] D. Slepian et J. Wolf., "Noise less coding of correlated information sources," IEEE Trans. Inform. Theory, Vol.19, No.4, PP. 471–480, 1973.
- [2] A. Wyner et J.Ziv. "The rate-distortion function for source coding with side information at the decoder," IEEE Trans. Inform. Theory, Vol. 22, No.1, PP.1–10, 1976.
- [3] X. Chen, X. Zhang, L. Zhang et al., "A wireless capsule endoscope system with low power controlling and processing ASIC," IEEE Transactions on Biomedical Circuits and Systems, Vol.3, No.1, pp.11–22, 2009.
- [4] L. Xiaoying, J. Gu, Y. Xie, J. Xiong, et W. Qin, "A new approach to detecting ulcer and bleeding in Wireless capsule endoscopy images," In proc. International Conference on Biomedical and Health Informatics (BHI), IEEE-EMBS, pp.737-740, 2012.
- [5] S. A. Karkanis, D. K. Iakovidis, D. E. Maroulis, D. A. Karras et M. Tzivras, "Computer-Aided Tumor Detection in Endoscopic Video Using Color Wavelet Features," IEEE Transactions on Information Technology in Biomedicine, Vol.7, No.3, pp.141-152, 2003.
- [6] C. E. Shannon., A mathematical theory of communication. Bell SystemTech.J, 27: 379–423 623–656, 1948.
- [7] C.E.Shannon et R.G.Gallager, "lower bounds to error probability for coding on discreet memoryless channels," information and control, 10, 65- 102, 1967.
- [8] R. Zamir, S. Shamai, and U. Erez, "Nested linear/lattice codes for structured multiterminal binning," IEEE Transactions on Information Theory, vol. 48, no. 6, pp.1250-1276, June 2002.
- [9] A. Wyner, "Recent results in the Shannon theory," IEEE Transactions on Information Theory, Vol. 20, No. 1, pp. 2-10, January 1974.
- [10] Y. Zhao and J. Garcia-Frias, "Data compression of correlated non-binary sources using punctured turbo codes," in IEEE Data Compression Conference, DCC 2002,Snowbird, UT, April 2002.
- [11] P. Mitran and J. Bajscy, "Coding for the Wyner-Ziv problem with turbo-like codes,"in IEEE International Symposium on Information Theory, ISIT 2002, Lausanne, Switzerland, June 2002.
- [12]D. Varodayan, A. Aaron, et B. Girod, "Rate-adaptive codes for distributed source coding," Signal Processing, vol. 86, no. 11, pp. 3123–3130, November 2006.

- [13] A. Liveris, Z. Xiong, and C. Georghiades, "Compression of binary sources with side information at the decoder using LDPC codes," *IEEE Communication Letters*, vol. 6, no. 10, pp. 440-442, October 2002.
- [14] Q. Xu, V. Stankovic, and Z. Xiong, "Layered Wyner-Ziv video coding for transmission over unreliable channels," *Signal Processing*, vol. 86, no. 11, pp. 3212-3225, November 2006.
- [15] S. Pradhan and K. Ramchandran, "Distributed source coding using syndromes (DISCUS): Design and construction," *IEEE Transactions on Information Theory*, vol. 49, pp. 626-643, March 2003.
- [16] S. Yaman et G. AlRegib, "A low-complexity video encoder with decoder motion estimator," *ICASSP2004, CANADA*, May 2004.
- [17] Gary J. Sullivan et al, "Overview of the high efficiency video coding (HEVC) standard," *IEEE transaction circuits and systems for video technology*, Vol.22, No.12, Dec 2012.
- [18] A. Aaron, R. Zhang et B. Girod, "Wyner-Ziv Coding of Motion Video," *Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, USA, November 2002.
- [19] R. Puri et K. Ramchandran, "PRISM: A New Robust Video Coding Architecture Based on Distributed Compression Principles," *40th Allerton Conference on Communication, Control and Computing*, Allerton, USA, October 2002.
- [20] J. Chou, S. Pradhan, et K. Ramchandran, "Turbo and trellis-based constructions for source coding with side information," in *IEEE Data Compression Conference, DCC 2003*, Snowbird, UT, March 2003.
- [21] Y. Yang, S. Cheng, Z. Xiong, et W. Zhao, "Wyner-Ziv coding based on TCQ and LDPC codes," in *Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, November 2003.
- [22] B. Girod, A. Aaron, S. Rane, et D. Rebollo-Monedero, "Distributed video coding," *Proceedings of the IEEE*, Vol. 93, No. 1, pp. 71-83, January 2005.
- [23] M. Ouaret, F. Dufaux, et T. Ebrahimi, "Codec-independent scalable distributed video coding," in *IEEE International Conference on Image Processing, ICIP 2007*, San Antonio, TX, September 2007.
- [24] M. Quaret, F. Dufaux, et T. Ebrahimi, "Fusion-based multiview video coding," in *ACM International Workshop on Video Surveillance and Sensor Networks, VSSN 2006*, Santa Barbara, CA, October 2006.

- [25] X. Guo, Y. Lu, F. Wu, D. Zhao, et W. Gao, "Wyner-Ziv-based multiview video coding," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 18, No. 6, PP. 713-724, June 2008.
- [26] R. Puri, A. Majumdar et K. Ramchandran, "PRISM: A Video Coding Paradigm with Motion Estimation at the Decoder", *IEEE Trans. on Image Process.*, Vol. 16, No. 10, PP. 2436 - 2448, Oct. 2007.
- [27] A. Aaron, S. Rane, E. Setton et B. Girod, "Transform-domain Wyner-Ziv Codec for Video", *VCIP'04*, San Jose, California, USA, January 2004.
- [28] X. Artigas, J. Ascenso, M. Dalai, S. Klomp, D. Kubasov and M. Oualet, "The DISCOVER Codec: Architecture, Techniques and Evaluation", *PCS'07*, Lisbon, Portugal, November 2007.
- [29] DISCOVER Page, <http://www.img.lx.it.pt/~discover/home.html>.
- [30] J. Slowack, J. Skorupa, S. Mys, P. Lambert, C. Grecos, and R. Van de Walle, "Flexible distribution of complexity by hybrid predictive-distributed video coding," *Signal Processing: Image Communication*, Vol. 25, No. 2, PP. 94–110, February 2010.
- [31] J. Chou, S. Pradhan, et K. Ramchandran, "Turbo and trellis-based constructions for source coding with side information," in *IEEE Data Compression Conference, DCC 2003*, Snowbird, UT, March 2003.
- [32] Y. Yang, S. Cheng, Z. Xiong, et W. Zhao, "Wyner-Ziv coding based on TCQ and LDPC codes," in *Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, November 2003.
- [33] S. Cheng et Z. Xiong, "Successive refinement for the Wyner-Ziv problem and layered code design," *IEEE Transactions on Signal Processing*, Vol. 53, No. 8, PP. 3269-3281, August 2005.
- [34] T. Wiegand, G. J. Sullivan, G. Bjøntegaard, et A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 13, No. 7, PP. 560-576, July 2003.
- [35] C.-Y. Chen, S.-Y. Chien, Y.-W. Huang, T.-C. Chen, T.-C. Wang, et L.-G. Chen, "Analysis and architecture design of variable block-size motion estimation for H.264/AVC," *IEEE Transactions on Circuits and Systems-I: Regular Papers*, Vol. 53, No.2, PP. 578-592, February 2006.
- [36] ITU-T, "ITU-T Recommendation H.263: Video coding for low bit rate communication," ITU-T, 1995.

- [37] C. Brites, J. Ascenso, et F. Pereira, "Improving transform domain Wyner-Ziv video coding performance," in IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2006, Toulouse, France, May 2006.
- [38] D. Kubasov, K. Lajnef, et C. Guillemot, "A hybrid encoder/decoder rate control for Wyner-Ziv video coding with a feedback channel," in IEEE MultiMedia Signal Processing Workshop, MMSP 2007, Chania, Crete, October 2007.
- [39] J. Ascenso, C. Brites, et F. Pereira, "Improving frame interpolation with spatial motion smoothing for pixel domain distributed video coding," in 5th EURASIP Conference on Speech and Image Processing, Multimedia Communications and Services, Smolenice, Slovak Republic, June-July 2005.
- [40] J. Ascenso, C. Brites, et F. Pereira, "Content adaptive Wyner-Ziv video coding driven by motion activity," in IEEE International Conference on Image Processing, ICIP 2006, Atlanta, GA, October 2006.
- [41] S. Klomp, Y. Vatis, et J. Ostermann, "Side information interpolation with sub-pel motion compensation for Wyner-Ziv decoder," in International Conference on Signal Processing and Multimedia Applications, SIGMAP 2006, Setúbal, Portugal, August, 2006.
- [42] D. Kubasov, J. Nayak, et C. Guillemot, "Optimal reconstruction in Wyner-Ziv video coding with multiple side information," in IEEE MultiMedia Signal Processing Workshop, MMSP 2007, Chania, Crete, October 2007.
- [43] Said Benierbah, Mohammed Khamadja, "Generalized Hybrid Intra and Wyner-Ziv Video Coding", IEEE Transactions on circuits and systems for video technology, vol. 21, no. 12 pp. 1929-1934, December 2011.
- [44] J. Ascenso, C. Brites, and F. Pereira, "Improving frame interpolation with spatial motion smoothing for pixel domain distributed video coding," in Proceedings of the 5th EURASIP Conference on Speech and Image Processing, Multimedia Communications and Services, Smolenice, Slovak Republic, July 2005.
- [45] C. Brites, J. Ascenso, and F. Pereira, "Improving transform domain Wyner-Ziv video coding performance," in Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'06), vol. 2, pp. 525–528, Toulouse, France, May 2006.
- [46] K. Misra, S. Karande, and H. Radha, "Multi-hypothesis based distributed video coding using LDPC codes," in *Proceedings of Allerton Conference on Commun, Control and Computing*, Allerton, Ill, USA, September 2005.

- [47] L. Wei, Y. Zhao, and A. Wang, "Improved side-information in distributed video coding," in *Proceedings of International Conference on Innovative Computing, Information and Control*, Beijing, China, August-September 2006.
- [48] J. Ascenso, C. Brites, and F. Pereira, "Motion compensated refinement for low complexity pixel based distributed video coding," in *Proceedings of IEEE Conference on Advanced Video and Signal Based Surveillance (AVSS '05)*, pp. 593–598, Como, Italy, September 2005.
- [49] X. Artigas and L. Torres, "Iterative generation of motion compensated side information for distributed video coding," in *Proceedings of IEEE International Conference on Image Processing (ICIP '05)*, pp. 833–836, Genova, Italy, September 2005.
- [50] Nikos Deligiannis et al, "Side-information-dependent correlation channel estimation in hash-based distributed video coding", *IEEE Transactions on Image Processing*, vol.21, no.4, pp. 1934-1949, 2012.
- [51] Huynh Van Luong et al, "Re-estimation of motion and reconstruction for distributed video coding," *IEEE Transactions on Image Processing*, vol.29, no.7, pp.2804-2819, July 2014.
- [52] E. Martinian, A. Vetro, J. S. Yedidia, J. Ascenso, A. Khisti, D. Malioutov, "Hybrid distributed video coding using codes," in *Proceedings of the 8th IEEE Workshop on Multimedia Signal Processing (MMSp '06)*, pp. 258–261, Victoria, Canada, October 2006.
- [53] M. Maitre, C. Guillemot, and L. Morin, "3-D model-based frame interpolation for distributed video coding of static scenes," *IEEE Transactions on Image Processing*, vol. 16, pp. 1246–1257, 2007.
- [54] N. Deligiannis et al., "Maximum likelihood Laplacian correlation channel estimation in layered Wyner-Ziv coding," *IEEE Transactions on signal processing*, Vol.62, No.4, February 15, 2014.
- [55] Brites C, Pereira F, "Correlation noise modeling for efficient pixel and transform domain Wyner-Ziv video coding," *IEEE Transactions on Circuits and Systems for Video Technology* 2008,18(9):1177-1190.
- [56] Wang H, Cheung NM, Ortega A, "A framework for adaptive scalable video coding using Wyner-Ziv techniques," *EURASIP Journal on Applied Signal Processing* 2006, **2006**:-18.

- [57] Vatis Y, Klomp S, Ostermann J, “Enhanced reconstruction of the quantised transform coefficients for Wyner-Ziv coding,” *Proceedings of IEEE International Conference on Multimedia and Expo (ICME '07), July 2007, Beijing, China 172-175.*
- [58] Kubasov D, Nayak J, Guillemot C, “Optimal reconstruction in Wyner-Ziv video coding with multiple side information,” *Proceedings of the 9th IEEE International Workshop on Multimedia Signal Processing (MMSP '07), October 2007, Crete, Greece 183-186.*
- [59] Weerakkody WARJ, Fernando WAC, Kondo AM, “An enhanced reconstruction algorithm for unidirectional distributed video coding,” *Proceedings of the 12th IEEE International Symposium on Consumer Electronics (ISCE '08), April 2008, Algarve, Portugal.*
- [60] Yun-Chung Shen et al, “Efficient Realtime Distributed Video Coding by Parallel Progressive Side Information Regeneration,” *IEEE SENSORS Journal*, Vol.17, No. 6, pp.1872-1883, January 2017.
- [61] L. Abid, “Épidémiologie des cancers en Algérie : problématique des registres des cancers”, *African Journal of Cancer*, vol.1, No.2,pp-98-103, 2009.
- [62] G.Iddan, G.Meron, et al., “Wireless capsule endoscopy,” pp. 405–417, *Nature* (2000).
- [63] T.Nakamura and A.Terano, “Capsule endoscopy: past, present, and future”, *Journal of Gastroenterology*, Vol.43, No.2, pp.93-99, February 2008.
- [64] Ben-Soussan E, Savoye G, Antonietti M, et al. “Is a 2-liter PEG preparation useful before capsule endoscopy?,” *J Clin Gastroenterol*, Vol. 39, No. 38, pp.1-4, 2005.
- [65] Dai N, Gubler C, Hengstler P, et al. “Improved capsule endoscopy after bowel preparation,” *Gastro intest Endosc*; vol. 61, pp.28-31, 2005.
- [66] Sharma V K, Eliakim R, Sharma P, et al. “ICCE consensus for esophageal capsule endoscopy,” *Endoscopy*; Vol.37, 2005.
- [67] Spada C, Hassan C, Marmo R, et al. “Meta-analysis shows colon capsule endoscopy is effective in detecting colorectal polyps,” *Clin Gastroenterol Hepatol*, Vol. 8, No.6, pp.516-522, 2010.
- [68] Spada C, Hassan C, Munoz-Navas M, et al. “Second-generation colon capsule endoscopy compared with colonoscopy,” *Gastrointest Endosc*, Vol.43, No. 3, pp. 581-589, 2011.
- [69] Eliakim R, Yassin K, Niv Y, et al. “Prospective multicenter performance evaluation of the second-generation colon capsule compared with colonoscopy,” *Endoscopy*, Vol.41, pp.1026-1031, 2009

- [70] Adler DG, Chand B, Conway JD, et al. "Capsule endoscopy of the colon," *Gastrointest Endosc*; vol.68, pp.621-623, 2008.
- [71] P. Swain, "At a watershed? Technical developments in wireless capsule endoscopy", *J Dig Dis.*, Vol. 11, No. 5, pp. 259-65, 2010.
- [72] T. Yokoi, et al, "Capsule endoscope", U.S. Patent 7 505 802, Mar. 17, 2009.
- [73] O.Y. Mang, et al., "Multiple LEDs luminous system in capsule endoscope", in *Proceedings SPIE the International Society for Optical Engineering*, vol. 8, no. 7, Feb. 6 2008.
- [74] H. Yu, G.S. Irby, D.M. Peterson, M.-T. Nguyen, G. Flores, N. Euliano and R. Bashirullah, "A Printed Capsule Antenna for Medication Compliance Monitoring", *IEEE Electronics Letters*, Vol. 43, No. 22, pp. 1179–1181, 2007.
- [75] S. Kwak, K. Chang, et Y. J. Yoon, "Ultra-wide band spiral shaped small antenna for the biomedical telemetry," in *Proc. Asia-PacificMicrow.*, Vol. 1, pp. 241-244, Conf., 2005.
- [76] H.H. Wenzl, K.D. Fine, L.R. Schiller, J.S. Fordtran, "Determinants of decreased fecal consistency in patients with diarrhea," *Gastroenterology*, Vol. 108, No. 6, pp. 1729-1738, 1995.
- [77] M. Mylonaki, A. Fritscher-Ravens, et P. Swain, "Wireless capsule endoscopy: a comparison with push enteroscopy in patients with gastroscopy and colonoscopy negative gastrointestinal bleeding," *Gut*, Vol.52, No.8, pp.1122–1126, 2003.
- [78] K. R. Canlas, B. M. Dobozi, S. Lin et al., "Using capsule endoscopy to identify GI tract lesions in cirrhotic patients with portal hypertension and chronic anemia," *Journal of Clinical Gastroenterology*, Vol.42, No.7, pp.844–848, 2008.
- [79] G. C. Sturniolo, V. DiLeo, M. G. Vettorato et al., "Small bowel exploration by wireless capsule endoscopy: results from 314 procedures," *American Journal of Medicine*, Vol.119, No.4, pp.341–347, 2006.
- [80] G. Pan, G. Yan, W. Xin, and J. Cheng, "A video wireless capsule endoscopy system powered wirelessly: design, analysis and experiment," *Measurement Science and Technology*, Vol.22, pp.1–9, 2011.
- [81] D. Turgis et R. Puers, "Image compression in video radio transmission for capsule endoscopy," *Sensors and Actuators A*, Vol.123-124, pp.129–136, 2005.
- [82] X. Li, M. Zhang, Z. Wang, et A. Bermak, "Smart image sensor with integrated low complexity image processing for wireless endoscope capsules," *Tsinghua Science and Technology*, Vol.14, No.5, pp.586–592, 2009.

- [83] J. Keller, C. Fibbe, F. Volkeetal, "Remote magnetic control of a wireless capsule endoscope in the esophagus safe and feasible: results of a randomized, clinical trial in healthy volunteers," *Gastrointestinal Endoscopy*, Vol.72, No.5, pp.941–946, 2010.
- [84] E. Morita, N. Ohtsuka, Y. Shindoetal, "In vivo trial of a driving system for a self-propelling capsule endoscope using a magnetic field (with video)," *Gastrointestinal Endoscopy*, Vol.72, No.4, pp.836–840, 2010.
- [85] P. Swain, A. Toor, F. Volkeetal, "Remote magnetic manipulation of a wireless capsule endoscope in the esophagus and stomach of humans (with videos)," *Gastrointestinal Endoscopy*, Vol.71, No.7, pp.1290–1293, 2010.
- [86] L. Stankovic, V. Stankovic, et S. Cheng, "Distributed compression: Overview of current and emerging multimedia applications," in *IEEE International Conference on Image Processing, ICIP 2011, Brussels, Belgium, September 2011*.
- [87] D.E. Boudechiche, S. Benierbah et M. Khamadja, "A new approach to generate side information for distributed video coding in capsule endoscopy," In *6th Int. Symposium On Communication, Control and Signal processing (ISCCSP), Greece, 124-127, 2014*
- [88] D.E. Boudechiche, "Distributed Video Coding of the Chroma Components for the Capsule Endoscopy images," in *proc. of Networking and Advanced Systems, Annaba-Algeria 2015, pp.86-89*.
- [89] X. Chen, X. Zhang, L. Zhang et al., "A wireless capsule endoscope system with low power controlling and processing ASIC," *IEEE Transactions on Biomedical Circuits and Systems*, Vol. 3, No. 1, pp. 11–22, 2009.
- [90] X. Xiang, et al "A low-power digital IC design inside the wireless endoscopic capsule," *IEEE Journal of Solid-State Circuits*, Vol. 41, No. 11, pp. 2390–2400, 2006.
- [91] K. Wahid, S. B. Ko, and D. Teng, "Efficient hardware implementation of an image compressor for wireless capsule endoscopy applications," in *Proceedings of the International Joint Conference on Neural Networks*, pp. 2761–2765, June 2008.
- [92] P. Turcza et M. Duplaga, "Low-power image compression for wireless capsule endoscopy," in *Proceedings of the IEEE International Workshop on Imaging Systems and Techniques*, pp. 1–4, May 2007.
- [93] M. Lin, L. R. Dung, et P. K. Weng, "An ultra-lowpower image compressor for capsule endoscope," *BioMedical Engineering Online*, Vol. 5, No. 14, pp. 1–8, 2006.

- [94] N. Bourbakis, G. Giakos, et A. Karargyris, "Design of new-generation robotic capsules for therapeutic and diagnostic endoscopy," in Proceedings of the IEEE International Conference on Imaging Systems and Techniques, pp.1–6, 2010.
- [95] M. Quirini, A. Menciassi, C. Stefanini, S. Gorini, G. Pernorio, et P. Dario, "Development of a legged capsule for the gastrointestinal tract: an experimental setup," in Proceedings of the IEEE International Conference on Robotics and Biomimetics, pp.161–167, July 2005.
- [96] S. Hosseini et M. B. Khamesee, "Design and control of a magnetically driven capsule-robot for endoscopy and drug delivery," in Proceedings of the IEEE Toronto International Conference Science and Technology for Humanity, pp.697–702, September 2009.
- [97] D. O. Faigel et D. R. Cave, "Capsule Endoscopy," Saunders Elsevier, Amsterdam, The Netherlands, 2008.
- [98] N. Deligiannis et al., "Wyner-Ziv video coding for wireless lightweight multimedia applications," EURASIP Journal on Wireless Communications and Networking, pp. 106-126. 2012.
- [99] T. H. Khan et K. A. Wahid, "Low power and low complexity compressor for video capsule endoscopy," IEEE Transactions on Circuits and Systems for Video Technology, Vol.21, No.10, pp.1534-1546, 2011.
- [100] S. Golomb, "Run-length encodings," IEEE Transactions on Information Theory, Vol.12, No.3, pp.399–401, 1966.
- [101] S. Chung, T. Richardson, et R. Urbanke, "Analysis of sum-product decoding of low-density parity-check codes using a Gaussian approximation," IEEE Transactions on Information Theory, Vol. 47, No. 2, pp. 657-670, February 2001.
- [102] S.-Y. Chung, G. D. Forney, T. J. Richardson, and R. Urbanke, "On the design of low-density parity-check codes within 0.0045 db of the Shannon limit," IEEE Communications Letters, Vol. 5, No. 2, pp. 58-60, February 2001.
- [103] J. Abouei, J. D. Brown, K. N. Plataniotis, et S. Pasupathy, "Energy efficiency and reliability in wireless biomedical implant systems," IEEE Transactions on Information Technology in Biomedicine, Vol. 15, No. 3, pp. 456-466, May 2011.
- [104] P. D. Bradley, "An ultra low power, high performance medical implant communication system (MICS) transceiver for implantable devices," in IEEE Biomedical Circuits and Systems Conference, BioCAS 2006, London, UK, November-December 2006.
- [105] William B. Pennebaker, Joan L. Mitchell, "JPEG: Still Image Data Compression Standard", Springer Science & Business Media, 1991.

- [106] Thomas G. Lane 1991, 1992, 1993 of the Independent JPEG Group's.
- [107] R. Sidhu, *et al.*, “Gastrointestinal capsule endoscopy: from tertiary centres to primary care.” *BMJ*. Vol. 332, pp.528–531, 2006.
- [108] G. Bjontegaard, “Calculation of average PSNR differences between RD-curves,” ITU-T Video Coding Experts Group (VCEG), TX, Document VCEG-M33, Apr. 2001.
- [109] Boudechiche, Djamel Eddine, Said Benierbah, and Mohammed Khamadja. “Distributed video coding based on vector quantization: Application to capsule endoscopy,” *Journal of Visual Communication and Image Representation*, Vol.49, pp.14-26, November 2017.
- [110] Lloud (S.P), “Least square quantization in PCM,” *IEEE trans. On Information Theory*, Vol. 28, No.2, mars 1982, pp.129-137.
- [111] R.Gray, “Vector Quantization,” *IEEE ASSP Magazine*, Vol.1, No.2, PP.4-29, April 1984.
- [112] T.C.Lu, C.Y. Chang, “A survey of VQ codebook generation,” *J. Inf. Hiding multimed. Signal proc. (JIHMSP)* Vol.1, No.3, pp.190-203, 2010.
- [113] Linde, Y. Buzo, A et Gray, R. M, “An algorithm for vector quantizer design,” *IEEE Trans. Commun.*, Vol.28, pp.84-95, January 1980.
- [114] A. K. Jain, et C. R. Dubes, “Algorithms for clustering data,” Prentice Hall, 1988.
- [115] R. Martins, C. Brites, J. Ascenso et F. Pereira, “Adaptive deblocking filter for transform domain Wyner–Ziv video coding,” *IET Image Process.*, Spec. Issue Distributed Video Coding, Vol.3, No.6, pp.315–328, Dec. 2009.
- [116] C. Brites and F. Pereira, “Correlation noise modeling for efficient pixel and transform domain wyner-ziv video coding,” *IEEE transactions on circuits and systems video technology*, Vol.18, No.9, pp.1177-1190, September 2008.
- [117] D. Kubasov, J. Nayak and C. Guillemot, “Optimal Reconstruction in Wyner-Ziv Video Coding with Multiple Side Information,” in *Proc. Of IEEE International Workshop on Multimedia Signal Processing (MMSP)*, Crete. Greece, pp.183-186, 2007.
- [118] C. Brites, F. Pereira, “Distributed video coding: Assessing the HEVC upgrade,” *Signal Processing: Image Communication*, Vol.32, No.81–105, 2015.

- [119] N. Deligiannis, “Distributed Video Coding for Wireless Lightweight Multimedia Applications,” Ph.D. dissertation, Dept. Elect. Info., Brussel Univ., June 2012.
- [120] C. Brites et F. Pereira, “Correlation noise modeling for efficient pixel and transform domain wyner-ziv video coding,” *IEEE transactions on circuits and systems video technology*, Vol.18, No.9, pp.1177-1190, September 2008.
- [121] Roberto M. E. Filho et al., “Codebook Calibration Method for Vector Quantizers Implemented at the Focal Plane of CMOS Imagers,” *IEEE Transactions on Circuits and Systems for Video Technology*, 2015.
- [122] C-Y Chen, S-Y Chien, Y-W Huang, T-C Chen, T-C Wang, L-G Chen, “Analysis and architecture design of variable block-size motion estimation for H.264/ AVC,” *IEEE Trans Circ Syst I*. Vol.53, No.2, pp.578–592, 2006.
- [123] Jan Hanca, Nikos Deligiannis, Adrian Munteanu, “Real-Time Distributed Video Coding for 1K-pixel Visual Sensor Networks, *Journal of Electronic Imaging*,” *Journal of Electronic Imaging* Vol. 25, No.4, 2016.
- [124] M. H. Taieb, J.Y. Chouinard et D. Wang, “Low Complexity Rate Estimators for Low Latency Wyner-Ziv Video Decoders,” *Engineering Letters*, Vol.21, No.1, pp.1-9, Jan. 2013.
- [125] Y.-S. Pai, Y.-C. Shen, and J. Wu, “High efficient distributed video coding with parallelized design for ldpca decoding on cuda based gpgpu,” *Journal of Visual Communication and Image Representation*, Vol.23, No.1, pp.63–74, Jan 2012.
- [126] CAO Yue, GAO Shaoshuai, ZHANG Can et QIU Gengfeng, “Towards practical distributed video coding for energy-constrained networks,” *Chin J Electron*, Vol.25, No.1, pp.121–130.
- [127] <http://niremf.ifac.cnr.it/tissprop/htmlclie/htmlclie.php>.

- [128] Hsin-Ping Yang et al, “An Improved Distributed Video Coding with Low-Complexity Motion Estimation at Encoder,” IEEE International System-on-Chip Conference (SOCC), 2015.
- [129] Toshiyuki Nozawa et al, “A Parallel Vector-Quantization Processor Eliminating Redundant Calculations for Real-Time Motion Picture Compression,” IEEE JOURNAL OF SOLID-STATE CIRCUITS, Vol.35, No.11, Nov 2000.
- [130] T Brack, M Alles, T Lehnigk-Emden, F Kienle, N When, NE L’Insalata, F Rossi, M Rovini, L Fanucci, “Low complexity LDPC code decoders for next generation standards,” Design, Automation, and Test in Europe, France, pp.1–6, April 2007.
- [131] P. Cosman, et al., “Evaluating Quality of Compressed Medical Images: SNR, Subjective Rating, and Diagnostic Accuracy,” *Proc. of the IEEE*, Vol. 82, No.6, pp.919-932, June 1994.
- [132] L.-R.Dung, Y.-Y.Wu, H.-C.Lai, and P.-K.Weng, “A modified h.264 intra-frame video encoder for capsule endoscope,” IEEE Biomedical Circuits Systems Conf., Baltimore, MD, 2008.
- [133] J.Berens, M. Mackiewicz, and G. D. Bell, “Stomach Intestine and Colon Tissue discrimination for Wireless Capsule Endoscopy Images,” Proceedings of SPIE, pp.283-290, 2005.
- [134] A. Menciassi, G. Ciuti, and C. Cavallotti, “Future Developments of Video Capsule Endoscopy: Hardware,” In Video Capsule Endoscopy, Springer Berlin Heidelberg, pp.543-556, 2014.
- [135] G.K. Wallace, “The JPEG still picture compression standard,” *Commun. ACM* Vol.34 pp.30–44, 1991.