N° d'ordre : Année : 2016/2017

République Algérienne Démocratique et Populaire
Ministère de l'enseignement supérieur et de la recherche scientifique

UNIVERSITÉ DJILLALI LIABES DE SIDI BEL ABBES
Faculté de Génie Electrique
Département de Télécommunications

THESE DE DOCTORAT

Pour l'obtention du Diplôme de Docteur
Filière : Electronique
Spécialité: Télécommunications

Titre de la thèse

Contribution à la protection des images et vidéos contre les erreurs de transmission, pertes de paquets et effacement: un nouveau schéma de compression multi-description utilisant la NSCT et OMP.

Présentée par :

M^elle^ **Amina NAIMI**

Le 02 Février 2016

Devant le jury composé de :

Président :
- *Mr. CHOUAKRI Sid Ahmed* — *Professeur à UDL-SBA*

Examinateurs :
- *Mr. SEDDIKI Ali* — *MCA à UDL-SBA*
- *Mme. BENAMRANE Nacéra* — *Professeur à USTO-MB-Oran*
- *Mr. BOUZIANI Merahi* — *Professeur au CU d'El-Bayadh*

Directeur de thèse :
- *Mr. BELLOULATA Kamel* — *Professeur à UDL-SBA*

Laboratoire de Télécommunications et Traitement Numérique du Signal.

Dédicaces

À tous ceux qui me tiennent à cœur

mes très chères parents.

mes frères et sœurs.

Tous mes amis.

Résumé

Un signal transmis est soumis à de nombreuses perturbations internes et externes au canal de transmission. Des techniques de traitement du signal, de codage et de modulation ont été développées ces dernières années pour améliorer la robustesse des canaux vis-à-vis du bruit. Néanmoins, le bruit n'est pas la seule source de perturbations. Pour optimiser les performances d'un canal de communication, la tâche est de trouver des solutions en terme de codage pour diminuer la probabilité d'erreurs sur les symboles reçus. Par conséquent, le codage conjoint source-canal peut être avantageux aux erreurs de transmission.

Dans cette thèse on propose un nouveau codeur vidéo MD. Nos méthodes se basent sur ce type de codage, la transformée NSCT, l'algorithme OMP. Ensuite un schéma de pré et post traitement basé sur le sur-échantillonnage spatiale est appliqué aux images de la séquence vidéo d'entrée. Dans ce schéma on a créé deux descriptions à transmettre à travers deux différents canaux. Les résultats d'expérimentations justifient les performances do nos méthodes MDC proposées utilisant la NSCT qu'avec la DWT dans les différents cas de perte d'une ou de plusieurs descriptions, ainsi dans le cas de perte de quelques trames à partir d'une ou plusieurs descriptions.

Mots clés : *MDC, NSCT, CT, DWT, OMP, pré et post traitement, H.264.*

Abstract

A transmitted signal is subjected to several internal and external disturbances due to communication channel. Signal processing techniques, coding and modulation have been developed in recent years to improve robustness against noise. However, the noise is not the only source of disturbance. To optimize the performance of a communication channel , the task is to find solutions in terms of coding to reduce the probability of received erroneous symbols. Therefore, the joint source-channel coding may be advantageous to free error transmission.

A new MD video coders for robust transmission over lossy networks are proposed. Our methods relies on MDC (Multiple Description Coding) scheme, NSCT (Non Subsampled Contourlet Transform), OMP (Orthogonal Matching Pursuit) algorithm for sparse representation of images in fact that NSCT is a redundant transform and a pre-post processing principle is applied to a video sequence in order to create two descriptions. The experimental results substantiate the performance of our proposed MD schemes using NSCT compared to DWT in cases of lose of some frames or completely one or more descriptions.

Keywords : *MDC, NSCT, CT, DWT, OMP, pre-post processing, H.264.*

ملخص

الإشارة المرسلة معرضة لمجموعة من الاضطرابات الداخلية والخارجية نتيجة لقناة الاتصال التي تغير نقاء الإشارة و يقصد بها الضوضاء. لقد تطورت تقنيات معاجلة الإشارات و التشفير(الترميز) وتعديل الإشارة خلال السنوات الأخيرة بغرض تحسين قوتها ضد الضوضاء والتداخلات. بالرغم من أن الضوضاء هي ليست السبب الوحيد للإضرابات التي تحدث للإشارة. لتحسين أداء قناة الاتصال يجب إيجاد حل عن طريق التشفير حيث ذلك يؤدى إلى تقليل احتمالية حدوث أخطاء في الرموز الواردة إذ اقترح الباحثون برنامجا جديدا للترميز عن طريق الوصف المتعدد MDC. مخطط ترميزالفيديوالمقترح في المذكرة عن طريق الوصف المتعدد يستخدم مع الشبكات التي تعرف بفقدها للإشارة لتقوية الإرسال. طريقة عمل هذا المخطط تعتمد على ما يعرف بنظام ترميز أو تشفير الوصف المتعدد و التحويل الذي لا يجزئ العينة NSCT و مواصلة التطابق المتعامد OMP و هي خوارزميات للتمثيل المفرق و المضغوط للصور بسبب التحويل التكراري الذي لا يجزئ العينة. وقد قدمنا مخططا آخر عن طريق الوصف المتعدد بأسلوب معالجة يطبق على الفيديو قبل و بعد الإرسال لكي ينشئ وصفين للإشارة الأصلية .نتائج التجارب أثبتت الأداء الأمثل لمقترحنا المقدم باستخدام الوصف المتعدد للإشارة و التحويل الذي لا يجزئ العينة بالمقارنة مع التحويل بالموجات المنفصلة DWT في حالة فقدان أطر البيانات أو فقدان وصف كلي واحد أو أكثر.

الكلمات المفتاحية : الوصف المتعدد, التحويل الذي لا يجزئ العينة, التحويل بالموجات المنفصلة, المعالجة قبل و بعد الإرسال, المشفرH264.

Remerciements

Je tiens tout d'abord à remercier l'Éternel Dieu vivant, par qui j'ai eu le souffle nécessaire pour la rédaction de ce document et qui m'a fait savoir que la science est la vraie connaissance de sa majesté.

Je Remercie chaleureusement tous les membres de ma famille pour leurs conseils avisés, leurs aides multiples et leur patience.

Merci à Mon directeur de thèse Professeur Kamel BELLOULATA, qu'il trouve ici ma sincère reconnaissance pour ses compétences, sa patience, et que j'ai beaucoup appris de lui, tant d'un point de vu scientifique et humain.

Je voudrai ensuite remercier tous les membres de jury :

- *Mr. CHOUAKRI Sid Ahmed, Professeur à l'université Djillali Liabès de Sidi Bel Abbes.*
- *Mr. SEDDIKI Ali, Maitre de conférence "A" à l'université Djillali Liabès de Sidi Bel Abbes.*
- *Mme. BENAMRANE Nacéra, Professeur à l'université Université des sciences et de la technologie d'Oran - Mohamed-Boudiaf.*
- *Mr. BOUZIANI Merahi, Professeur et Recteur du centre universitaire Nour El Bachir d'Elbayadh.*

Qui m'ont honoré avec leur acceptation de juger mon travail et pour l'intérêt porté à cette thèse.

Je remercie tous les membres et camarades du laboratoire de Télécommunications et Traitement Numérique du Signal (TTNS) pour les bons moments passés, accompagnés de leurs conseils.

Amina NAIMI

Chapitre II : Codage et compression des images et vidéo

ARQ	Automatic Repeat request.
FEC	Forward Error Correction.
MDC	Multiple Description Coding.
CT	Transformée en Contourlets.
NSCT	Non Subsampled Contourlet Transform.
OMP	Orthogonal Matching Pursuit.
DWT	Discret Wavelet Transform.
TFD	Transformé de Fourier Discrète.
TCD	Transformée en Cosinus Discrète.
TH	Transformée de Hadamard.
JPEG	Joint Photographic Experts Groupe.
FFT	Fast Fourier Transform.
SFQ	Space frequency Quantization.
AMR	Analyse Multirésolution.
LP	Laplacian Pyramid.
DFB	Directionnal Filter Bank.
NSP	Non Subsampled Pyramid.
NSFB	Nonsubsampled Filter Bank.
NSDFB	Nonsubsampled Directionnal Filter Bank.
RVB	Rouge, vert, Bleu.
AVC	Advanced Video Coding.
CCD	Charge Couple Device.
BMP	Windows BitMap.
TIF	Tagged Image File Format.
GIF	Graphics Interchange Format.
ppp	points par pouce.
dpi	dots per inch.
DCT	Discret Cosine transform.
MPEG	Moving Picture Experts Groupe.
MIT	Massachusetts Institute of Technology.

RLC	Run Length Coding.
DPCM	Differential Pulse Code Modulation.
QV	Quantification Vectorielles.
ISO	International Organization for Standardization.
IEC	International Electrotechnical Commission.
AC	Alternative Component.
DC	Direct Component.
CCITT	International Telephone and Telegraph Consultative Committee.
JVT	Joint video Team.
ITU	International Telecommunication Union.
UVLC	Unified Variable Lenght Coding.
CAVLC	Contest Adaptative Variable Lenght Coding.
CABAC	Context Adaptative Binary Arithmetic Coding.
QP	Quatification Parameter.
SD	Single Description.
MDSQ	Multiple Description Scalar Quantization.
MDLVQ	Multiple Description Lattice Vector Quantization.
LOT	Lapped Orthogonal Transform.
MDTC	Multiple Description Transform Coding.
IDCT	Inverse Discret Cosine Transform.
PSNR	Peak Signal to Noise Ratio.
MV	Motion Vector.
BP	Basis Pursuit.
MP	Matching Pursuit.
StOMP	Stagewise Orthogonal Matching Pursuit.
CoSaMP	compressive sampling matching pursuit.
ROMP	Regularized Orthogonal Matching Pursuit.
BPD	Basis Pursuit Denoising.
PCT	Pairwise Correlating Transform.
EQM	Erreur Quadratique Moyenne.

SPM	Self Projected Matching.
DDWT	Dual Tree Wavelet Transform.
NS	Noise Shaping.
QCIF	Quart Common Intermediate Format.
SR	Sparsity Ratio.
MCDI	Multimedia Content Description Interface.
HD	Haute Définition.
HEVC	High Efficiency Video Coding.
VCEG	Video Coding Experts Group.

Cette thèse décrit les résultats de recherche que nous avons mené au laboratoire *TTNS* (Télécommunications et Traitement Numérique du Signal) de l'Université Djilali Liabes de Sidi Bel Abbes. Ces recherches ont été dirigées par Pr BELLOULATA Kamel, professeur à l'université Djillali Liabes au département de télécommunications. Cette thèse porte sur le codage à descriptions multiples des images et séquences animées.

Depuis l'avènement de 1$ère$ numérique, la compression devient alors une nécessité pour faciliter la transmission des données sur des réseaux à perte de paquets. Les algorithmes de compression effectifs recherchent les redondances d'un objet pour les éliminer et ainsi diminuer la taille de sa représentation. L'image est utilisée dans différents domaines tels que la médecine, la robotique, la télévision, la vidéoconférence, la vidéosurveillance, le Web et d'autres. Ce besoin important de l'image dans le quotidien a introduit de nouvelles problématiques parmi lesquelles nous pouvons citer la compression d'images, l'extraction d'objets, la segmentation, etc.

L'intérêt grandissant du monde des télécommunications et de ses consommateurs pour le réseau Internet a vu de nouveaux types de services apparaitre sur ce réseau ces dernières années. Ceux ci incluent la transmission de données multimédia en temps réel, comme le son, la vidéo, les images de synthèse animées, etc.

L'utilisation des réseaux sans fils ou l'internet implique que la transmission de données sans erreurs peut être atteinte par la retransmission des paquets perdues ou endommagés par des mécanismes de control d'erreurs tel que la Requête Automatique de Répétition (*ARQ* : Automatic Repeat request) qui assure une transmission sans erreurs. Ces techniques sont révélées être très efficaces pour la transmission vidéo sans fils. La retransmission des trames corrompues introduit un délai supplémentaire qui pourrait être inacceptable pour la communication en temps réel. La correction d'erreur directe (*FEC* : Forward Error Correction) implique l'ajout de données redondantes au signal compressé, ce qui permet au décodeur de corriger les erreurs à un certain niveau [1]. Les techniques de contrôle d'erreur *ARQ*, *FEC* ou même les deux ensemble ne peuvent pas être facilement adaptées aux émissions en temps réel. Par conséquent, le codage conjoint source-canal est souvent préféré.

En particulier, *MDC* (Multiple Description Coding) est avérée être un moyen efficace pour fournir une résistance aux erreurs de transmission. En effet, Le codage à descriptions multiples consiste à représenter l'information originale en plusieurs unités redondantes et indépendantes appelées descriptions.

Dans notre travail, on s'intéresse à un schéma de codage *MD* basé sur des transformées redondantes tout en respectant l'objectif de la compression des données. Ce manuscrit s'organise en quatre chapitres :

Chapitre I, consiste à décrire les nouvelles transformations directionnelles. La compression de données peut être réalisée par transformation. La transformée en ondelettes est un outil largement utilisé en traitement du signal et d'image, l'une de ses applications principales est la compression d'images du fait sa capacité à compacter l'énergie sur un petit nombre de coefficients permettant un codage efficace de l'image. Elle correspond à la première étape de la chaîne de compression classique.

L'analyse multirésolution est l'un des outils les plus utilisés dans l'analyse et le traitement d'images numériques du fait qu'elle permet d'analyser le signal dans différents niveaux de résolutions. L'analyse multirésolution est capable d'analyser et interpréter les structures grâce au fait qu'elle décompose l'image dans plusieurs échelles de résolutions et d'orientations différentes. Ces propriétés ont rendu l'analyse multirésolution un outil puissant pour faire la compression, la segmentation, le codage, la transmission de l'image, etc. Parmi les domaines sur lesquels nous pouvons appliquer ces outils, nous citons le codage.

Chapitre II, Dans ce chapitre on s'intéresse à donner un rappel sur l'image et la vidéo numérique. Les principaux objectifs dans l'élaboration des équipements vidéo numériques étaient de réduire les besoins en bande passante et en espace de stockage tout en proposant un niveau de qualité visuelle au moins égal à celui des équipements analogiques. Puis, on passe à présenter un bref aperçu sur quelques standards de compression de l'image et la vidéo.

Chapitre III, Un état de l'art sur le codage à descriptions multiples (*MDC*) et l'approximation parcimonieuse des signaux. Dans le cas à description unique, la perte d'un paquet bloque la reconstruction et nécessite sa retransmission, dans le cas descriptions multiples la perte de paquet est moins grave. La réception d'un nombre suffisant de paquets permet de reconstituer une partie du signal initial. Il s'agit donc d'une forme de codage conjoint. Le *MDC* prend le parti de répartir l'information différemment.

En effet, le codeur fournit non pas une seule mais plusieurs chaines de bits $D_1, D_2 \ldots D_N$ appelées descriptions, Chacune d'entre elles constitue une unité indépendamment transmissible sur différents canaux.

On a consacré une partie à l'état de l'art méthodologique sur l'approximation parcimonieuse des signaux. Les méthodes de représentation adaptative, alliant redondance et parcimonie, fournissent cette représentation efficace. Nous décrivons alors les approches de poursuite de base, poursuite adaptative et poursuite adaptative orthogonale.

Chapitre IV, Se consacre à la présentation de nos méthodes de codage *MDC* qui se base sur des décompositions en sous bandes multirésolution *CT* et *NSCT* d'une image et de la vidéo et une comparaison avec des schémas *MDC* existants. Dans une première partie, Nous présentons une méthode *MDC* sur l'image fixe. On s'occupe en premier temps à la comparaison entre trois types de transformées : *DWT*, *CT* et *NSCT*. Nous appliquons ensuite l'algorithme d'approximation de poursuite adaptative orthogonale *OMP* sur les sous bandes de la *NSCT* afin de réduire le nombre de coefficients à coder. Une deuxième partie du chapitre est consacrée à l'étude des performances de la transformée *NSCT* dans un schéma de codage vidéo à descriptions multiples basé sur le principe de pré et poste traitement spatiale des images de la séquence vidéo. Nous présentons au cours du dernier chapitre les expérimentations et l'évaluation de nos méthodes proposées.

Le manuscrit s'achève par une conclusion générale des différents travaux menés au cours de cette thèse.

I.1. Introduction :

Les bases orthogonales discrètes ont fait la preuve de leur utilité depuis des décennies aussi bien pour les mathématiciens que les ingénieurs en proposant des représentations vectorielles efficaces. En particulier, les transformées linéaires sont au centre de nombreuses méthodes en traitement du signal et des images. Elles sont employées par exemple dans l'analyse, la détection, le filtrage, le débruitage ou la compression de données, comme la transformée de Fourier discrète (*TFD*), la transformée en cosinus discrète (*TCD*) ou la transformée de Hadamard (*TH*). Les transformées en ondelettes telles qu'on les connait aujourd'hui ont pris corps dans les années 80. Cet outil permet de représenter un signal à différentes résolutions ce qui se révèle utile pour des applications aussi bien en analyse qu'en débruitage ou en compression, des codeurs basés sur les transformations en ondelettes, comme ceux de type JPEG2000, ont été proposés. Pour plus d'informations concernant la construction des bases d'ondelettes nous renvoyons au ouvrage [2]. Notons que l'ouvrage [3] regroupe une grande partie des articles fondateurs sur la théorie des ondelettes. Dans le domaine du traitement d'image, la notion d'analyse multirésolution apparait souvent. Elle incorpore et unifie des techniques développées pour différentes disciplines dont le codage en sous-bandes de signaux. Durant les dernières décennies, plusieurs chercheurs ont travaillé sur les transformées multirésolution pour remédier à la limite majeure de la transformée de Fourier (*FFT*) qui ne garde que l'information fréquentielle et perd l'information spatiale de l'image. Les transformées multirésolution permettent, non seulement l'interprétation de l'information fréquentielle, mais aussi sa localisation dans l'image. Les transformées multirésolution visent essentiellement à assurer les caractéristiques suivantes :

- ***La multirésolution*** : La représentation devrait permettre aux images d'être successivement approximées, d'une résolution grossière à une résolution fine.

- ***La localisation*** : Les éléments de base dans la représentation doivent être localisés dans le domaine spatial et fréquentiel.

- ***L'échantillonnage critique*** : Pour certaines applications (la compression), la représentation devrait constituer une base ou un cadre avec une redondance de petite taille.

- ***La directionnalité*** : La représentation doit contenir des éléments de base orientés à une variété de directions, beaucoup plus que les quelques directions qui sont offertes par les ondelettes séparables.

- ***L'anisotropie*** : Pour capturer des contours lisses dans les images, la transformée devrait avoir des éléments de base qui permettent de faire la recherche directionnelle de contours. Ceci est réalisé en utilisant, lors de la décomposition une variété de formes allongées avec différentes directions.

Dans ce chapitre, nous présentons trois types de transformées qui ont été utilisées fréquemment dans le domaine de traitement des images, à savoir, la transformée en ondelettes, la transformée en contourlets et sa variante redondante Non sous-échantillonnée. La seconde section de ce chapitre présente les transformées directionnelles. Nous introduisons dans la troisième section la représentation multirésolution des images. Nous finirons ce chapitre par une conclusion dans la quatrième section.

I.2. Les nouvelles transformées directionnelles :

La représentation efficace de l'information visuelle est au cœur de nombreux problèmes en traitement d'images incluant la compression, le débruitage, et l'extraction de primitives. Récemment, il est devenu évident que les transformations usuelles de type séparable, telle que la transformée en ondelettes, ne sont pas nécessairement les mieux adaptées pour la représentation des images qui semblent former une catégorie restreinte et limitée des possibilités de représentations des images et de signaux multidimensionnels. Cette limitation est due au fait que de telles représentations ne prennent pas en compte la régularité des structures géométriques d'une image. Par conséquent, plusieurs travaux récents ont mis l'accent sur la recherche de nouvelles techniques permettant une description de l'image.

Basées sur cette observation, d'autres montrent qu'il est possible de définir des représentations multi-échelle donnant naissance à de nouvelles transformées directionnelles plus adaptées à l'extraction de structures géométriques lisses et continues telles que les contours d'objets. Certaines reposent sur des bancs de filtres directionnels analysant l'image à des échelles, positions et orientations données, d'autres suivent une approche adaptative décrite par un modèle géométrique donnant explicitement la direction d'analyse locale.

Nous allons dans ce chapitre voir les transformées directionnelles existantes pour pallier aux problèmes induits par des transformées de type séparable tel que les ondelettes.

Parmi ces nouvelles représentations directionnelles cherchant à remédier aux problèmes induits par les ondelettes *2D* séparables, on peut trouver :

I.2.1. Transformée de Radon :

La transformée de Radon [4] consiste à projeter l'image sur un certain nombre d'orientations en intégrant l'image le long de la direction orthogonale à la projection, puis à réaliser la transformée de Fourier de ces projections. La reconstruction s'obtient en plaçant, pour chaque orientation de projection choisie, les coefficients de Fourier obtenus le long de cette même orientation, dans le domaine fréquentiel. On obtient l'image reconstruite en effectuant ensuite une transformée de Fourier *2D* inverse. La reconstruction parfaite pour cette transformée continue s'obtient pour un nombre de projections infini, parcourant l'ensemble des orientations possibles.

La transformée de Hough est un cas particulier de transformée de Radon pour une image à valeurs binaires, et s'utilise principalement pour la reconnaissance de formes. La transformée de Radon est très utilisée en tomographie, Une reconstruction approximative de l'image recherchée, d'autant plus précise que le nombre de directions de projection est élevé, est obtenue par transformée de Radon inverse. Notons que pour une image discrète carrée dont la taille est un nombre premier, la discrétisation proposée dans [5] peut s'appliquer pour obtenir une transformée de Radon discrète à reconstruction parfaite peu redondante.

I.2.2. Ridgelets :

Les ridgelets [6][7] forment une extension naturelle de la transformée de Radon pour un nombre limité de directions, en se basant sur des fonctions d'ondelettes pour contrôler la précision en orientation et garantir la reconstruction parfaite. Dans [8], une famille de ridgelets fenêtrées est proposée sous le nom d'orthoridgelets. Du fait du compromis sur l'incertitude temps-fréquence, ce fenêtrage implique une discrétisation en orientation pour une seule échelle donnée.

Cependant, en partant d'une fenêtre de la taille de l'image et en considérant toutes les familles d'orthorigdelets définies sur le partitionnement dyadique de cette fenêtre initiale en sous fenêtres, il est possible d'obtenir une décomposition en ridgelets multiéchelle. Notons qu'il existe également une transformée en ridgelets finie, proposée dans [7], en se basant sur la discrétisation des transformées de Radon finies [5]. Il est remarquable de préciser que cette transformée est non-redondante, nécessitant toutefois d'avoir une image carré dont la taille est un nombre premier. Diverses versions discrètes de la transformée en Ridgelets, menant à des implantations algorithmiques, ont été développées [9][10][11][12].

Malheureusement, par construction, la transformée en Ridgelets ne s'avère efficace que pour la caractérisation des contours rectilignes, ce qui nous amène à la transformée en Curvelets.

I.2.3. Curvelets :

Motivée par les limitations de la méthodologie des Ridgelets, la transformée en curvelet a été proposée dans [13]. Cette transformée se dérive des ridgelets multi-échelle. Une première décomposition permet d'obtenir une analyse multi-échelle en sous-bandes. Ces sous-bandes sont ensuite analysées par une transformée en orthoridgelets [13].

Une construction différente et plus générale, reposant sur la théorie des trames, est présentée dans [14]. Cette transformée a été utilisée avec succès dans le cadre du débruitage [15], elle a été utilisée également avec succès dans la fusion des images.

I.2.4. Beamlets :

La décomposition en beamlets [16] considère un partitionnement de l'image en quad-tree, puis effectue une transformée de Radon dans chaque bloc. Les coefficients de beamlets sont liés par une relation multi-échelle, où chaque beamlet à un niveau donné est décomposée en trois beamlets connexes au niveau suivant. Cette transformée permet d'approximer les courbes dans les images et d'en extraire les contours par sélection dans le graphe de connexité des beamlets.

I.2.5. Wedgelets :

La décomposition en wedgelets [17] représente une image par un quad-tree dans lequel chaque bloc est séparé en deux régions d'intensité différentes. Elle est donc, en quelque sorte, duale à la décomposition en beamlets en considérant les intégrales de l'image de chaque côté du segment représentant le contour plutôt que l'intégrale le long du segment. Bien que cette décomposition fournisse une approximation assez grossière des images, elle a été combinée dans [18] avec une décomposition en ondelettes séparables et un codeur *SFQ* [19] pour la compression. Cette transformée a été généralisée en dimension plus élevée sous le nom de surflets [20].

I.2.6. Bandelettes :

La transformée en bandelettes repose sur un modèle géométrique explicite de l'image pour effectuer une analyse orientée le long des contours. Dans une première approche [21], les contours sont représentés par des courbes paramétriques le long desquelles une ondelette séparable est déformée. Une seconde approche [22] considère un champ d'orientations plutôt qu'un nombre restreint de courbes. Le calcul des coefficients de bandelettes s'effectue de manière séparable en filtrant le long du champ d'orientation, à l'aide de techniques de lifting sur grilles irrégulières [23].

La complexité principale de cette transformée réside dans la recherche du champ d'orientation et l'optimisation débit-distorsion pour répartir le débit entre les différentes informations. Cette technique a également été adaptée à la compression de surfaces [24].

I.2.7 Directionlets :

Une transformée discrète à échantillonnage critique appelée transformée en directionlets est présentée dans [25][26]. Un filtrage séparable est effectué le long de deux directions d'analyse. Cette technique fournit une analyse anisotrope de l'image, bien adaptée aux images à forte structure géométrique. Notons que les ondelettes séparables peuvent être vues comme un cas particulier de directionlets.

I.3. Représentation Multirésolution des images :

Depuis le milieu des années 1970, de très nombreuses méthodes mettant en jeu des techniques liées aux bancs de filtres ont été développées. L'analyse multirésolution d'un signal revient à le décomposer à différentes échelles, en approximations et en détails. S. Mallat propose un algorithme rapide permettant de calculer les coefficients de détails et d'approximations en utilisant des filtrages et décimations successifs [27]. La généralisation de l'*AMR* (Analyse Multirésolution) aux signaux de plusieurs dimensions ne pose aucune difficulté d'ordre théorique si l'on utilise des ondelettes séparables.

I.3.1. Les Ondelettes :

La construction directe de bases discrètes orthonormées d'ondelettes *2D* repose sur la théorie de l'analyse multirésolution (*AMR*).

Cette analyse permet l'étude successive des approximations lisses d'un signal *2D* dans lesquelles les détails sont progressivement supprimées. L'un des éléments fondamentaux de l'analyse multirésolution *2D* est l'introduction d'une matrice de dilatation *d* qui définit le "processus" de lissage lors d'un changement de résolution. Cet outil permet de représenter un signal à différentes résolutions ce qui se révèle utile pour des applications aussi bien en analyse qu'en débruitage ou en compression, des codeurs basés sur les transformations en ondelettes, comme ceux de type *JPEG*2000, ont été proposés comme alternative à *JPEG* permettant de réduire les artéfacts produits par l'effet de bloc en codage d'image [28].

Pour plus d'informations concernant la construction des bases d'ondelettes nous renvoyons à l'ouvrage [2]. Une transformée en ondelettes est une analyse multirésolution d'un signal (ou d'une image). Autrement dit, le signal analysé est représenté à différentes échelles permettant de capturer des détails plus ou moins grossiers. Les différentes résolutions peuvent être vues comme des sous-espaces vectoriels $(V_j)_{j\in\mathbb{Z}} \subset L^2(\mathbb{R})$ emboités. L'analyse multirésolution d'un signal f consiste alors à réaliser des projections de f sur $(W_j)_{j\in\mathbb{Z}}$, les sous-espaces vectoriels complémentaires orthogonaux des $(V_j)_{j\in\mathbb{Z}}$.

Dans la pratique, une méthode d'implémentation des décompositions en ondelettes discrètes se fait suivant un schéma de bancs de filtres appliqués en cascade.

Le principe est d'appliquer d'abord un banc de filtres $[H_0, H_1]$ sur les données. Le résultat du filtrage par H_0 (qui sera supposé ici être un filtre passe bas) représente les coefficients d'approximation et le résultat par H_1 (supposé passe haut) représente les coefficients de détails du signal. On applique de nouveau ce banc de filtres sur le résultat du filtrage par H_0 et cette opération est répétée un nombre fini de fois.

À chaque étage de la décomposition en ondelettes on change en fait d'échelle dans la représentation du signal analysé. Si l'on sait exprimer un banc de filtres inverse de $[H_0, H_1]$, le schéma de synthèse se construit simplement de proche en proche, en utilisant une structure en cascade symétrique à celle utilisée pour l'analyse. La figure 1.2 présente également la partie de reconstruction. il est possible de travailler avec des transformées dites bi-orthogonales [29].

En pratique, pour calculer les coefficients d'approximations et de détails d'une image *I*, nous utilisons la généralisation de l'algorithme pyramidal présenté aux figures 1.1 et 1.2. On obtient pour un niveau de décomposition une sous image d'approximations $a_j(n,m)$et trois sous-images de détails : $d_j^1(n,m)$, $d_j^2(n,m)$, $d_j^3(n,m)$ selon l'orientation fréquentielle (horizontale, verticale et diagonale).

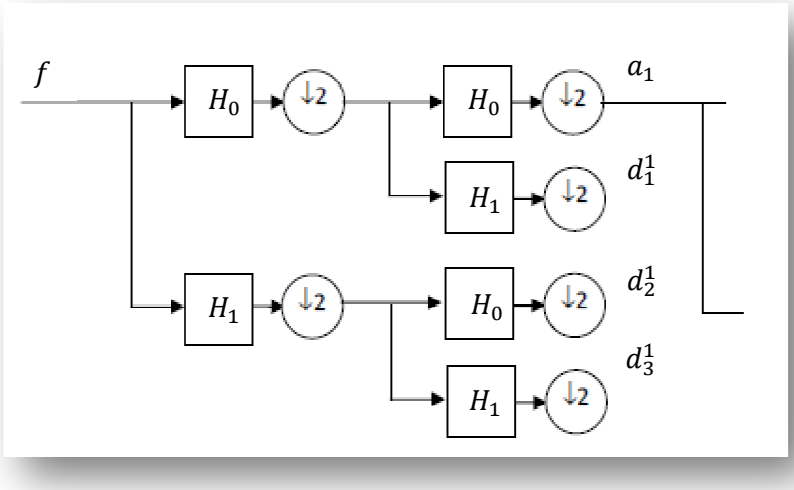


Figure 1.1. Algorithme pyramidal de décomposition d'une image.

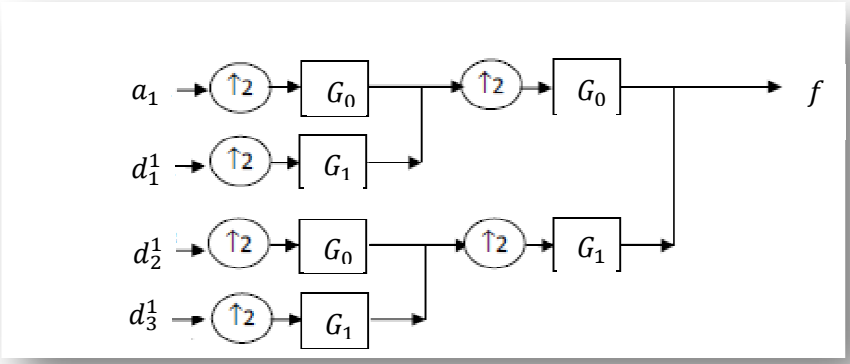
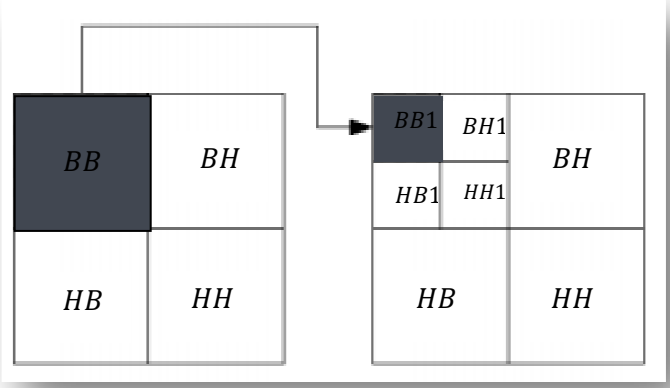
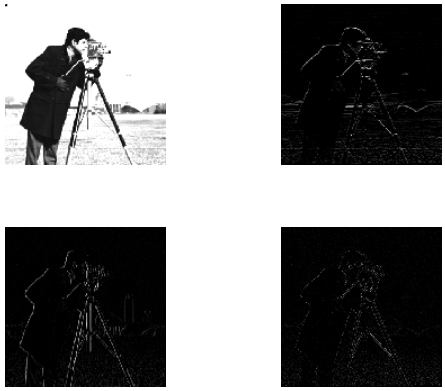


Figure 1.2. Algorithme pyramidal de reconstitution d'une image.



(a)



(b)

Figure 1.3. (a) Sous bandes de décomposition en ondelettes à deux niveaux
(b) Exemple de décomposition en ondelettes à deux niveaux**.**

La représentation en ondelettes amène à des algorithmes efficaces, en particulier elle conduit à des transformations rapides et des structures de données arborescentes.
Ce sont les principales raisons du succès des ondelettes dans de nombreuses applications de traitement et de communication du signal, par exemple, la

transformée en ondelettes a été adoptée comme la transformation de la nouvelle norme de compression d'images, *JPEG*2000 [30].

Les ondelettes en deux dimensions sont bonnes à isoler les discontinuités, mais ne capture pas les contours fins. En outre, les ondelettes séparables peuvent capturer seulement l'information directionnelle limitée, une caractéristique importante et unique de signaux multidimensionnels.

I.3.2. Transformée en Contourlets :

Plusieurs autres systèmes bien connus qui fournissent des représentations d'images directionnelles multi-échelle comprennent: ondelettes *2D* de Gabor [31], la transformée cortex [32], la pyramide orientables [33], les ondelettes directionnelles *2D* [34], brushlets [35], et les ondelettes complexes [36].

Les principales différences entre ces systèmes et la construction en contourlets est que les méthodes précédentes ne permettent pas un nombre différent de directions à chaque échelle, tout en réalisant un échantillonnage presque critique.

En outre, la représentation en contourlets utilise des bancs de filtres itératifs. Elle fournit une analyse multirésolution et directionnelle d'un signal bidimensionnel. La décomposition multirésolution est effectuée au moyen d'une pyramide laplacienne. L'analyse directionnelle s'effectue ensuite au moyen d'un banc de filtres directionnels à échantillonnage critique appliqué à chaque sous bande de la pyramide laplacienne excepté la sous-bande de basse fréquence.

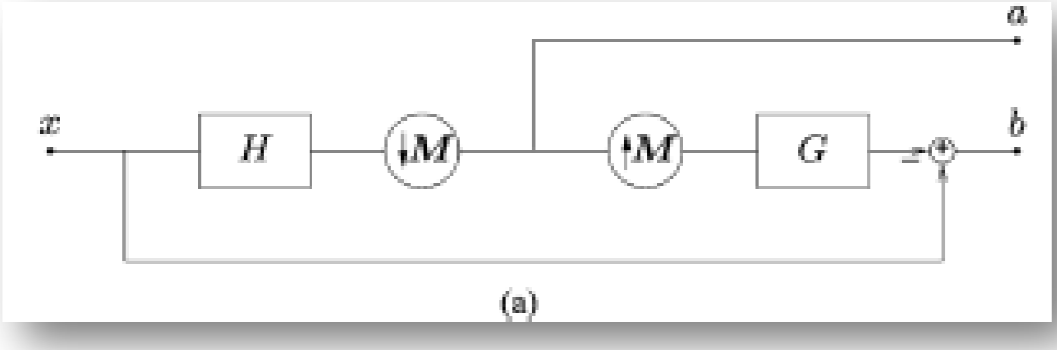
I.3.2.1 Pyramide Laplacienne :

Une seule façon d'obtenir une décomposition multi-échelle est d'utiliser la pyramide Laplacienne (*LP*) introduite par Burt et Adelson [37]. Á chaque niveau de décomposition, *LP* génère une version passe-bas sous-échantillonnée de l'originale et de la différence entre l'originale et la prédiction, résultant en une image passe-bande.
La figure 1.4.(a) représente ce processus de décomposition, où H et G sont appelés les filtres d'analyse et de synthèse passe-bas respectivement, et M

est la matrice d'échantillonnage. Le processus peut être répété sur le signal passe-bas sous-échantillonné.

L'inconvénient de la *LP* est le sur-échantillonnage implicite. Cependant, contrairement au régime des ondelettes critiquement échantillonné, la *LP* a la particularité que chaque niveau de la pyramide ne génère qu'une seule image passe-bande (aussi pour les cas multidimensionnels), et cette image n'a pas de fréquences "brouillé". Cette fréquence de brouillage se passe dans le banc de filtres d'ondelettes quand un canal passe-haut, après sous-échantillonnage, est replié dans la bande de fréquences basses, et par conséquent, son spectre est réfléchi. Dans la *LP*, cet effet est évité par sous-échantillonnage du canal passe-bas seulement.

Dans la théorie des trames et bancs de filtres sur-échantillonnés [39], ont montré que le système *LP* avec des filtres orthogonaux $h[n] = g[-n]$, dans ce cas, ils ont proposé l'utilisation de la reconstruction linéaire optimale en utilisant l'opérateur pseudo-inverse comme le montre la figure I.4.(b). La nouvelle reconstruction diffère des méthodes habituelles, où le signal est obtenu en ajoutant simplement la différence à la prédiction à partir du signal sous-échantillonnée, et ont montré une amélioration significative par rapport à la reconstruction d'habitude en présence de bruit.



(a)

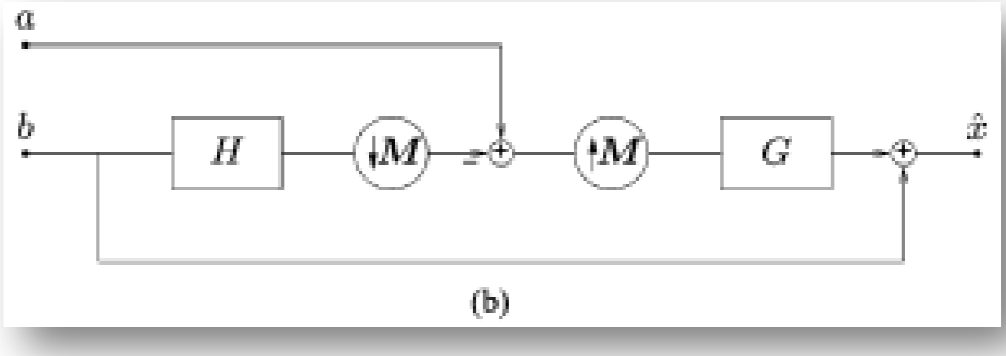


Figure 1.4. (a). Processus d'analyse *LP* (b). Processus de synthèse *LP* [43].

I.3.2.2. Banc de Filtre Directionnel

Bamberger& Smith [40] ont construit un banc de filtres directionnels *DFB* qui peut être décimé au maximum, tout en réalisant une reconstruction parfaite. Le DFB est efficacement mis en œuvre par l'intermédiaire d'une décomposition en arbre binaire de l −niveau qui conduit à 2^l sous-bandes avec un partionnement en fréquence cunéiforme comme représenté sur la figure 1.5. La construction originale du *DFB* [40] consiste à moduler l'image d'entrée et utilisant des bancs de filtres en quinconce avec des filtres en forme de diamant [41].

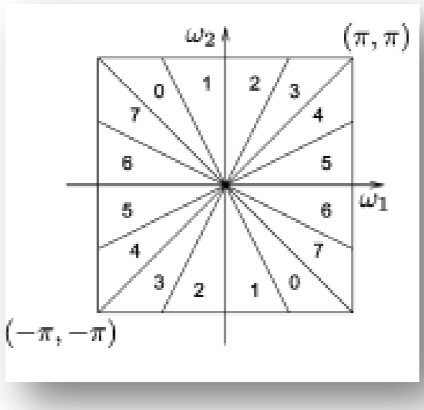


Figure 1.5. Partitionnement en fréquence à trois niveaux avec *DFB*[43].

M.N DO et al [42] ont proposé une nouvelle construction de *DFB* afin d'éviter la modulation de l'image d'entrée. leur *DFB* simplifié est intuitivement construit à partir de deux blocs. Le premier bloc est construit de deux canaux de banc de filtres quinconce [41] avec des filtres à ventilation (voir la figure 1.6.) qui divise un spectre *2D* en deux directions: horizontale et verticale. Le deuxième bloc de construction de *DFB* est un opérateur de cisaillement.

La figure 1.7. montre une application d'un opérateur de cisaillement, où une arête de direction $-45°$ devient verticale. Ainsi, la clé dans le *DFB* est d'utiliser une combinaison appropriée des opérateurs de cisaillement avec deux bancs de filtres en quinconce, afin d'obtenir un partitionnement en spectre *2D* désirée comme le montre la figure 1.5.

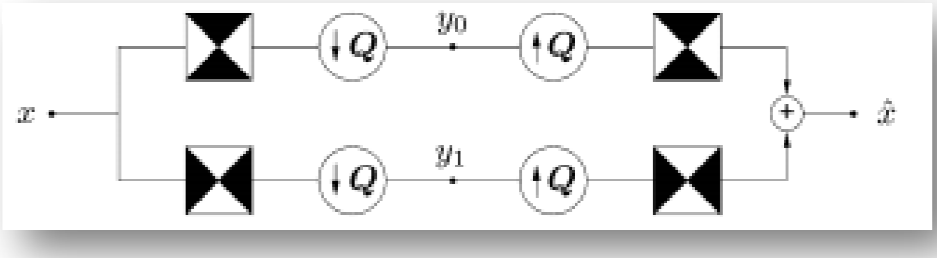


Figure 1.6. Banc de filtres en quinconce [43].

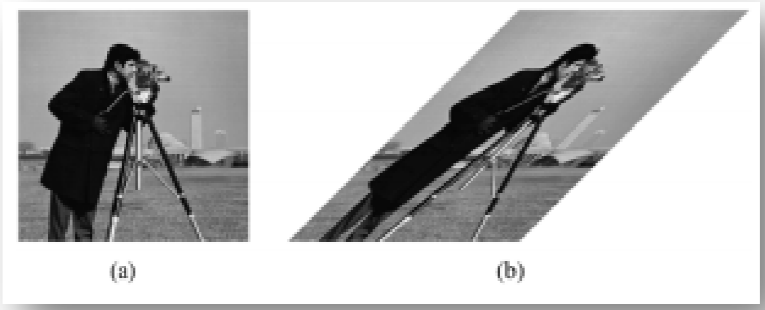


Figure 1.7. Exemple d'application d'un opérateur de cisaillement [43].

I.3.2.3. Banc de Filtre Directionnel Laplacien :

En combinant la *LP* et le *DFB*, nous sommes maintenant prêts à décrire leur combinaison dans une structure de banc de filtres double. Étant donné que le *DFB* a été conçu pour capturer la haute fréquence de l'image d'entrée, la basse fréquence est mal gérée.

En fait, avec le partitionnement fréquentiel représenté sur la figure 1.5, la basse fréquence aurait fuir en plusieurs sous-bandes directionnelles, par conséquent, le *DFB* seul ne fournit pas une représentation parcimonieuse pour des images. Ce fait fournit une autre raison de combiner le *DFB* avec une décomposition multi-échelle, où les basses fréquences de l'image d'entrée sont enlevées avant l'application du *DFB*.

Dans le nouveau banc de filtre pyramidal directionnel, La Pyramide Laplacienne (LP) est utilisée en premier pour capturer les discontinuités, suivie d'un banc de filtre directionnel (*DFB*) [43].

La figure 1.8. montre une décomposition multi-échelle et directionnelle en utilisant une combinaison d'une Pyramide Laplacienne (*LP*) et un Banc de Filtre Directionnel (*DFB*). Les images passe-bande à partir du *LP* sont introduites dans un *DFB* afin d'avoir l'information directionnelle. Le résultat combiné est une structure itérative à double bancs de filtres, nommée banc de filtre en Contourlets, qui décompose les images en sous-bandes directionnelles à des échelles multiples.

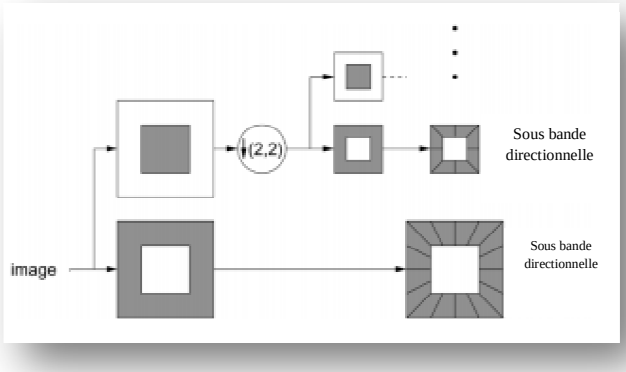


Figure 1.8. Diagramme en blocs de la transformée en Contourlets [43].

La figure 1.9 montre un exemple de décomposition du spectre fréquentiel par la transformée en contourlets.

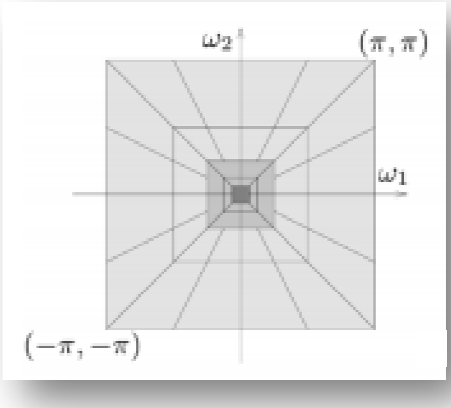


Figure 1.9. Partitionnement en fréquence avec *CT* [43].

La transformée en contourlets garantit une reconstruction parfaite de l'image. Son degré de redondance est relativement faible, les contours fins sont mieux représentés puisque des expérimentations ont déjà clairement montré que les contours lisses sont représentés de manière efficace par quelques coefficients contourlets localisés dans la bande à orientation appropriée [44][45].

La figure 1.10 illustre un exemple de décomposition de l'image Barbara utilisant la transformée en contourlets ou l'image est décomposée en deux niveaux pyramidaux, décomposés ensuite en quatre et huit sous bandes directionnelles. Les petits coefficients sont représentés en noir alors que les grands coefficients sont représentés en blanc.

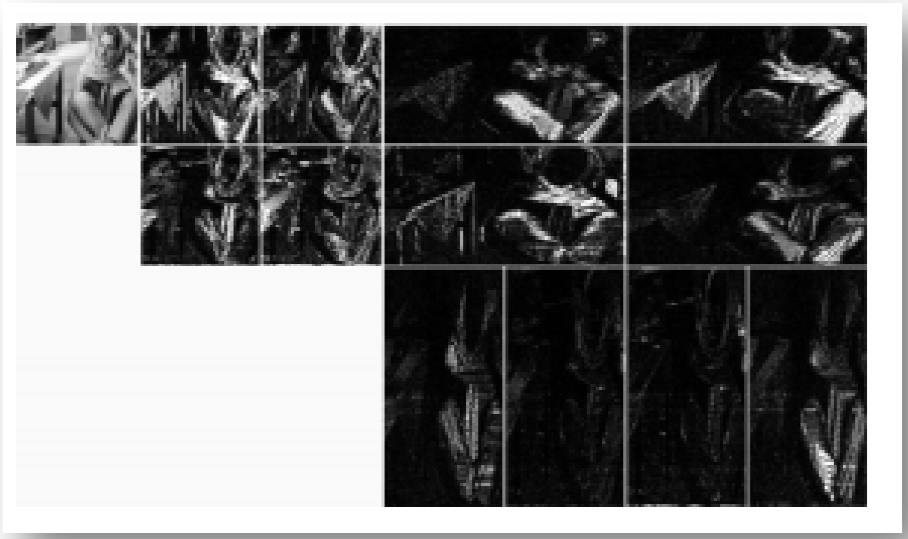


Figure 1.10. Exemple de décomposition multirésolution avec *CT* [43].

En raison de ce sur-échantillonnage et du sous-échantillonnage présents dans la pyramide laplacienne et les bancs de filtres directionnels, la transformée en contourlets n'est pas invariante en translation.

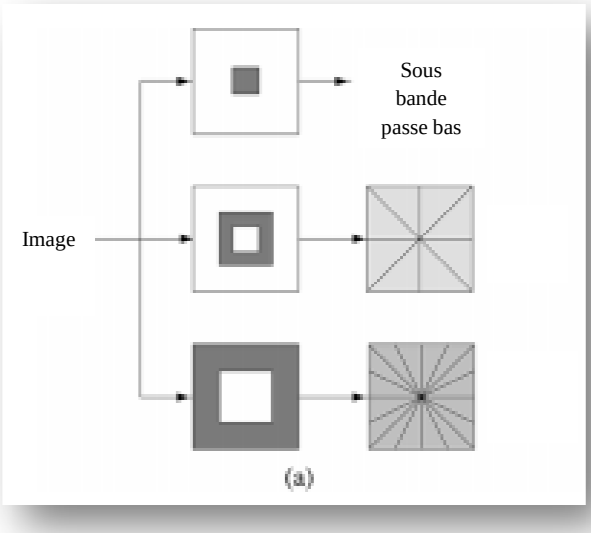
Un système est dit invariant par translation, si les coefficients de décomposition ou les sous-bandes résultantes sont de même taille. Cette propriété préserve mieux l'intégrité de l'image et l'est souhaitable dans les applications de traitement d'images, telles que la détection de contours, la fusion des images, etc.

Cette propriété n'est donc pas présente dans la transformée en contourlets à cause de la décimation à la sortie de chaque niveau de décomposition. De plus, le fait de considérer un banc de filtres fixe à échantillonnage critique pour l'analyse directionnelle implique de réaliser un compromis entre la précision sur la localisation spatiale des contours et la résolution en direction des filtres d'analyses. D'autre part, la sélectivité réduite introduit une réponse non négligeable des filtres directionnels en dehors de leur domaine fréquentiel idéal. Ceci se traduit en la présence d'aliasing indésirable dans les sous-bandes directionnelles. L'ensemble de ces points contribue à la dégradation des performances en fusion des images des contourlets.

I.3.3. Transformée en Contourlets Non Sous-échantillonnée :

la transformée en contourlets n'est pas invariante en translation, cela peut poser problème dans le cas par exemple de la détection de mouvements dans le domaine transformée. Un autre exemple qui illustre l'importance de l'invariance par translation est le débruitage par seuillage dans le domaine des ondelettes où le manque de l'invariance par translation fait apparaître le phénomène de Gibbs. Afin de pallier à ce problème, Cunha et al ont proposé une nouvelle transformée appelée transformée en Contourlets Non Sous-échantillonnée (Nonsubsampled Contourlet Transform, *NSCT*) [46].

La *NSCT* est entièrement invariante en translation, multi-échelle, et multidirectionnelle à une mise en œuvre rapide. La structure *NSCT* illustrée sur la figure 1.11.(a) consiste d'un banc de filtres qui partitionne le plan fréquentiel *2D* en sous-bandes, figure 1.11.(b). La transformée proposée peut être alors répartie en deux parties :



1) Une structure pyramidale non sous-échantillonnée qui assure la propriété multi-échelle.

2) Un Banc de filtres directionnelles non sous-échantillonnés qui assure la directionnalité.

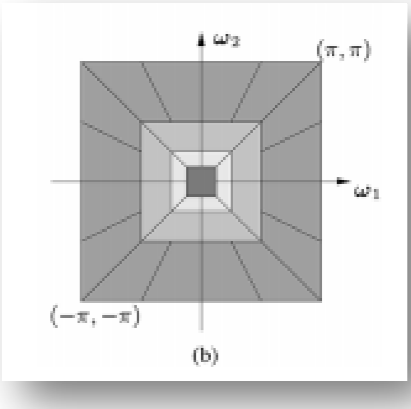


Figure 1.11. (a) Diagramme en blocs de la décomposition *NSCT*. (b) Partitionnement en fréquence [46].

I.3.3.1. Pyramide Non Sous-échantillonnée (NSP):

La propriété multi-échelle de la *NSCT* est obtenue à partir d'une structure de filtres invariante en translation, qui réalise une décomposition de sous-bande similaire à celle de la Pyramide Laplacienne. Ceci est atteint en utilisant deux canaux de bancs de filtres non sous-échantillonnés. La figure 1.12 illustre la décomposition pyramidale non sous-échantillonnée à trois étages, ces filtres donnent une analyse multi-résolution comme dans la figure 1.12 (b).

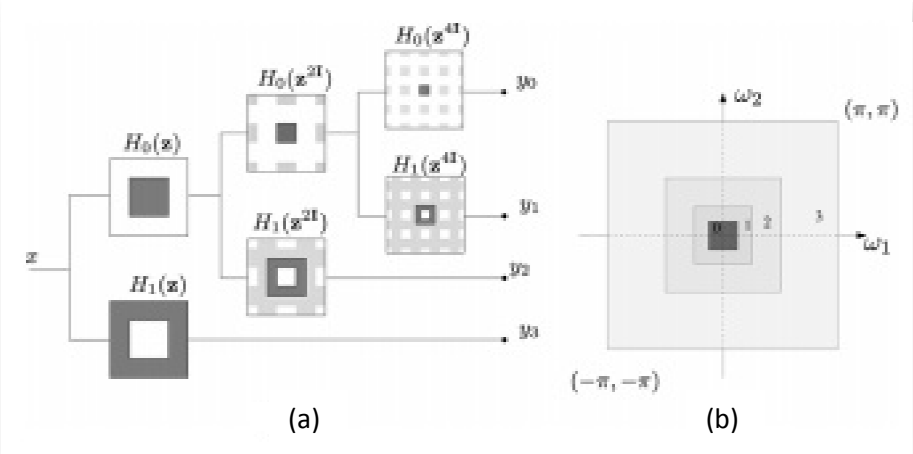


Figure 1.12 Trois étages de décomposition pyramidale [46].

Plus précisément, le *NSFB* (Nonsubsampled Filter Bank) est construit à partir de filtre passe-bas $H_0(z)$. On définit alors $H_1(z)$=1-$H_0(z)$, et les filtres de synthèse correspondants $G_0(z)$= $G_1(z)$=1. Une décomposition similaire peut être obtenue en retirant le sous-échantillonnage et sur-échantillonnage dans la pyramide laplacienne, puis échantillonnant les filtres en conséquence. $G_0(z)$ et $G_1(z)$ sont des filtres passe-bas et passe-haut respectivement. La reconstruction parfaite est atteinte ou les filtres satisfont la condition suivante :

$$H_0(z)\, G_0(z)+H_1(z)G_1(z) = 1 \qquad (1.1)$$

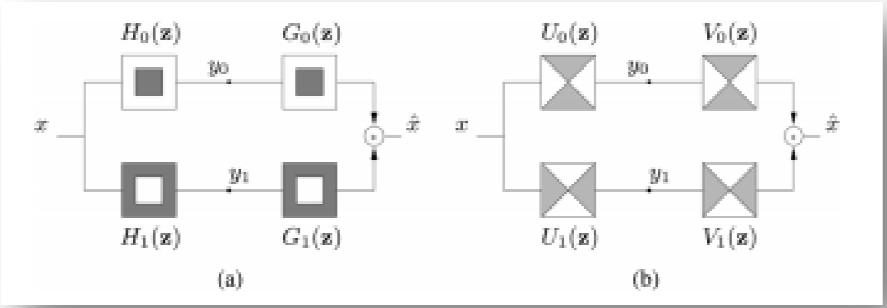


Figure 1.13. Réponse en fréquence (a) Pyramide non sous-échantillonnée (b) *DFB* non sous-échantillonée [46].

I.3.3.2. Banc de Filtre Directionnel Non Sous-échantillonné (NSDFB) :

Le banc de filtres directionnels de Bamberger et Smith [40] est construit en combinant deux canaux de bancs de filtres à ventilation critiquement échantillonnés et les opérations de ré-échantillonnage. Le résultat est un banc de filtres à structure arborescente qui partitionne le plan fréquentiel *2D* en quartiers directionnels.

Une version directionnelle invariante en translation est obtenue avec un *DFB* non sous-échantillonné (*NSDFB*). Le *NSDFB* est construit en éliminant le sous-échantillonnage et le sur-échantillonnage dans le *DFB* [47]. Cela se fait en alternant le sous-échantillonnage et le sur-échantillonnage dans chaque deux canaux de banc de filtres dans la structure arborescente *DFB* et échantillonnant les filtres en conséquence.

Il en résulte un arbre composé de deux canaux *NSFB*. La figure 1.14 illustre une décomposition à quatre canaux et la décomposition en fréquence directionnelle correspondante.

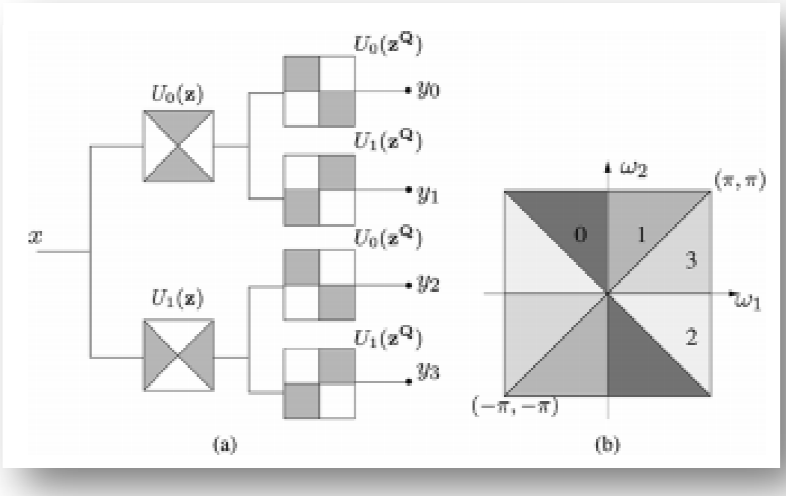


Figure 1.14. Partie Analyse d'un NSDFB [46].

I.3.3.3. Transformée en Coutourlets Non Sous-échantillonnée :

La structure *NSCT* est composée de deux canaux *NSFB* comme le montre la figure 1.13. La *NSCT* est ainsi construite en combinant la pyramide non sous-échantillonnée et les bancs de filtres directionnels non-sous-échantillonnés. La pyramide non sous-échantillonnée fournit une décomposition multi-échelle en divisant l'entrée en sous-bandes passe-bas et passe-haut, et les bancs de filtres directionnels non sous-échantillonnés fournissent une décomposition directionnelle en divisant la sous-bande passe-haut en plusieurs sous-bandes directionnelles. Ce schéma peut être répété sur la sous-bande passe bas à la sortie de la pyramide non sous-échantillonnée.

Comparée à la transformée en contourlets, il est plus facile et plus flexible de concevoir les bancs de filtres nécessaires qui mènent à une transformée en contourlets non sous-échantillonnée, avec une meilleure sélectivité et régularité fréquentielle. Ainsi, la NSCT est une décomposition d'image entièrement invariante par translation, multiéchelle, et multidirectionnelle avec une implémentation rapide. La *NSCT* s'est avérée être très efficace dans le débruitage et l'amélioration d'images comme il a été démontré dans [46].

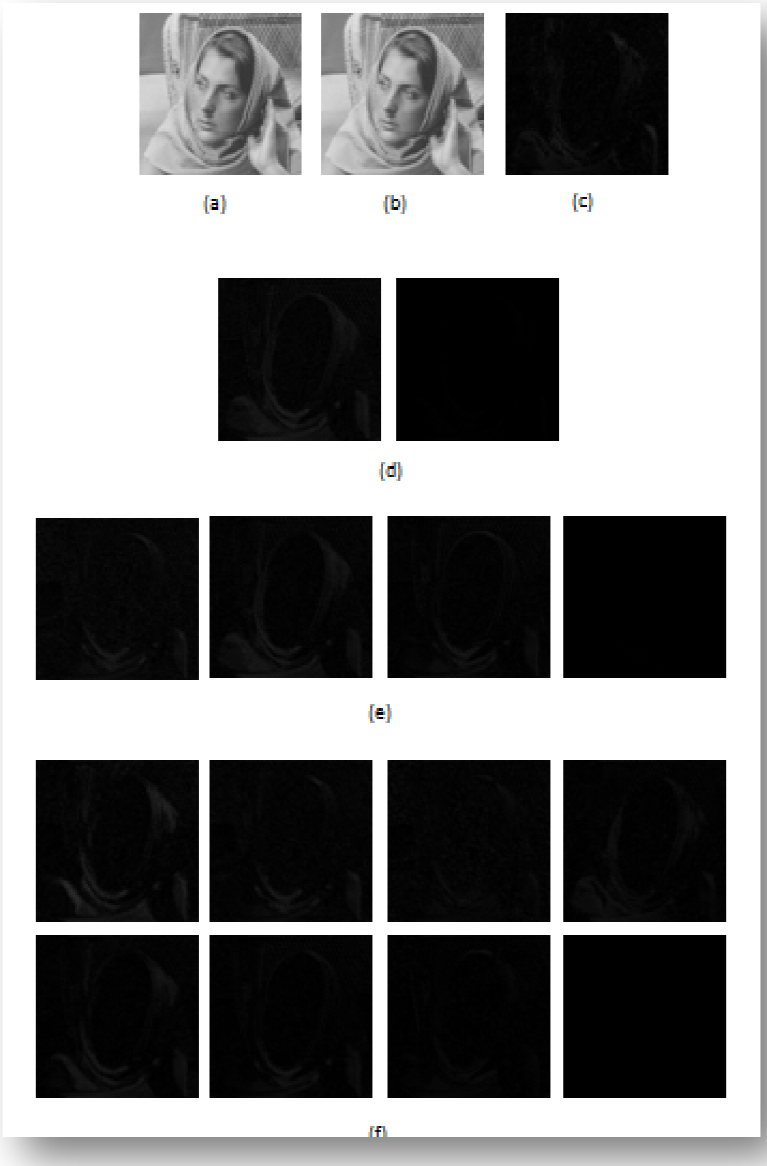


Figure 1.15. La transformée *NSCT* de l'image "Barbara".
(a) L'image originale. (b) Sous-bande passe-bas. (c) (d) (e) (f) Sous-bandes directionnelles passe-bande.

I.4. Conclusion

Les premières transformées directionnelles qui ont vu le jour sont les transformées de Radon et la transformée en Cortex, et ont été définies dans le domaine continu. Elles ont été source d'inspiration pour d'autres transformées, comme les Ridgelets et Curvelets, qui revient à utiliser une fonction d'ondelette pour discrétiser en orientations la transformée de Radon. Les contourlets s'inspirent en revanche du partitionnement fréquentiel proposé par la transformée en Cortex. La transformée en contourlets présente de plus l'intérêt notable d'avoir la meilleure sélectivité directionnelle. De plus, elle est peu redondante ce qui en fait une bonne candidate pour son application à la compression. La *NSCT* en revanche, offre à la fois une analyse discrète directionnelle et invariante par translation. Elle constitue ainsi une voie prometteuse pour des applications de traitement d'image telle que l'extraction de primitives. Une autre caractéristique importante de cette transformée est sa stabilité par rapport aux translations du signal d'entrée. Un exemple qui illustre l'importance de l'invariance par translation est le débruitage par seuillage dans le domaine des ondelettes où le manque de l'invariance par translation fait apparaître des artefacts de Gibbs autour des singularités dans le résultat du débruitage. Habituellement les transformées redondantes sont largement préférées pour des applications de traitement d'images telles que le débruitage, l'extraction de primitives et l'amélioration des images. En effet, les transformées redondantes sont plus flexibles et faciles à concevoir et peuvent surpasser de manière significative celles non redondantes. Les transformées d'images à échantillonnage critique en revanche, sont adoptées notamment pour des applications de codage et de compression.

Dans ce chapitre, nous avons passé en revue certaines nouvelles transformées. Ce sont des décompositions multi-échelle, qui opèrent selon une multitude d'orientations fréquentielles. En particulier, la transformée en contourlets non sous-échantillonnée, est une version discrète, donc adaptée aux images numériques et qui offre une meilleure sélectivité directionnelle possédant la propriété de l'invariance par translation. Ces caractéristiques intéressantes nous ont amené à l'exploitation de cette nouvelle transformée dans le codage et la compression des images et des séquences vidéo. Ceci fera l'objet des chapitres suivants.

II.1. Introduction :

Pour utiliser les images numériques, il faut comprimer les fichiers dans lesquels elles sont enregistrées. L'image consomme une quantité impressionnante d'octets quand elle est numérisée. Aujourd'hui, on parle de "qualité méga-pixel" pour les appareils photo numériques, cela signifie que chaque image comporte environ un million de pixels dont chaque pixel nécessite trois octets pour les composantes *RVB*. Les chercheurs ont remédié à ces contraintes pour réduire le volume global des images tout en conservant l'originale, un moyen consiste en la compression des images avec le minimum de dégradation et le maximum d'efficacité possible.

Les méthodes de compression et de codage réduisent le nombre de bits par pixel à stocker ou à transmettre, en exploitant la redondance informationnelle dans l'image. Le traitement d'image fait appel non pas à des images optiques classiques, mais à des images numériques. Le traitement d'image travaille sur des données chiffrées contenues dans l'image, modifient ces données qui sont ensuite utilisées pour construire une image transformée visualisable, d'archiver les images et de les transférer sur les réseaux informatiques. Les systèmes de compression d'image les plus performants à l'heure actuelle reposent sur trois étapes de traitement (figure 2.1) [48].

- L'étape de transformée consiste à appliquer une transformation généralement linéaire aux valeurs d'intensité représentant l'image afin d'en extraire les composantes principales. Parmi les approches les plus performantes dans ce domaine, les décompositions en ondelettes.

- L'étape de quantification dégrade généralement le signal. Les techniques de quantification utilisées dans les systèmes de compression d'image actuels sont généralement assez simples.

- Enfin, les codeurs de sous-bande utilisés en compression d'image reposent sur un codage progressif.

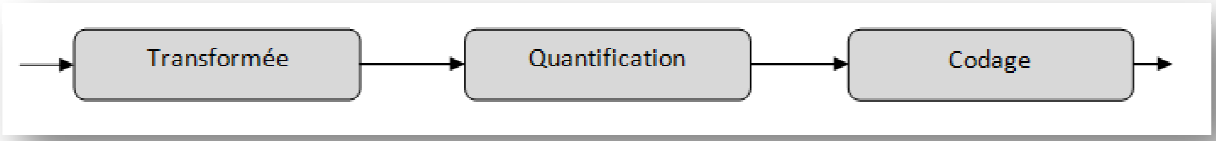


Figure 2.1. Système de compression par transformée.

Il existe des techniques de codage qui confèrent aux flux multimédia codés une certaine robustesse face aux éventuelles erreurs subies sur le réseau, la norme *H.264/AVC* décrite dans la dernière section de ce chapitre propose des méthodes pour augmenter la robustesse de la vidéo codée face aux pertes de paquets. Nous allons également décrire les principales techniques de codage présentes dans la norme *H.264/AVC*.

II.2. L'image numérique :

L'image numérique correspond à une matrice (ensemble ordonné à deux ou trois dimensions) de données numériques. Nous nous intéresserons uniquement aux images en deux dimensions. On peut concevoir ces images en *2D* comme un tableau de valeurs auxquelles on fait correspondre une position sur un plan (x, y) et une couleur pour visualiser l'image sur l'écran d'un ordinateur [47] [49].

Une image numérique *2D* est donc composée d'unités élémentaires appelées pixels "picture elements" qui représente chacune, une portion de l'image, codée par des valeurs numériques. Une image est définie par le nombre de pixels qui la composent en largeur et en hauteur et qui peut varier théoriquement presque à l'infini et l'étendue des teintes de gris ou des couleurs que peut prendre chaque pixel (dynamique de l'image) [47].

120	120	122	122	125	122	123
138	135	135	131	129	128	131
141	138	142	140	142	139	138
143	142	142	137	140	139	139
136	136	132	129	132	131	133
145	141	147	141	141	139	141
140	137	140	134	136	138	140
130	128	131	130	131	137	143
122	120	113	111	118	120	120
107	106	107	106	108	108	108
89	89	87	95	98	96	102
93	96	102	99	98	99	99
100	97	105	102	106	107	104
109	106	98	98	104	101	104
100	103	104	104	107	110	108

Figure 2.2. Détails des valeurs des pixels d'une image en niveaux de gris.

Toutes les données correspondant aux informations chiffrées contenues dans l'image sont structurées, afin de permettre leur stockage [48].

La numérisation d'une image est obtenue par l'intermédiaire d'un capteur et d'un numériseur, qui transforment un signal optique en un signal numérique. Le capteur est constitué par un ensemble de capteurs élémentaires, une barrette de *CCD* (Charge Couple Device), ça peut être une caméra, noir et blanc ou couleur, un appareil photo numérique ou un scanner. Le signal électrique est repris par un convertisseur analogique-numérique qui transforme les données continues en données numériques codées sur 1, 8, 16 ou 24 bits. Le codage utilisé définit le type de l'image (noir et blanc, niveaux de gris ou couleur) et sa profondeur. La taille des *CCD* et le pas d'échantillonnage de l'image au niveau de la carte de numérisation définissent la résolution spatiale de l'image. Le format de l'image est défini par l'entête du fichier et peut être précisé par son extension .bmp, .tif, .gif, .jpeg etc..., il renseigne sur le mode de présentation des données et leur degré de compression [47] [49].

II.2.1. notion de pixels :

Une image est constituée d'un ensemble de points appelés pixels. Le pixel représente ainsi le plus petit élément constitutif d'une image numérique. L'ensemble de ces pixels est contenu dans un tableau à deux dimensions constituant l'image [47] [49].

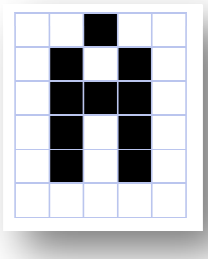


Figure 2.3. Pixel: le plus petit élément constitutif d'une image numérique.

Étant donné que l'écran effectue un balayage de gauche à droite et de haut en bas, on désigne généralement par les coordonnées [0,0] le pixel situé en haut à gauche de l'image, cela signifie que les axes de l'image sont orientés de la façon suivante [47]:

- L'axe x est orienté de gauche à droite.
- L'axe y est orienté de haut en bas, contrairement aux notations conventionnelles en mathématiques, où l'axe y est orienté vers le haut.

II.2.2. Résolution d'une image :

Elle est définie par le nombre de points image ou pixels représentant l'image, par unité de longueur de la structure à numériser, on exprime cette résolution en "pixels par pouce" (ppp) ou "dots per inch" (dpi). Un pouce représente $2.54\ cm$. Beaucoup de personnes expriment la résolution d'une image par son nombre total de pixels en x et y, mais c'est un abus de langage [47].

Plus le nombre de pixels est élevé par unité de longueur de la structure à numériser, plus la quantité d'information qui décrit cette structure est importante et plus la résolution est élevée. Lorsque la résolution diminue, la précision diminue puisque l'objet est représenté par un nombre moins important de pixels [47] [48].

La résolution permet ainsi d'établir le rapport entre le nombre de pixels d'une image et la taille réelle de sa représentation sur un support physique.

Une résolution de 300 dpi signifie donc 300 colonnes et 300 rangées de pixels sur un pouce carré ce qui donne donc 90000 pixels sur un pouce carré. La résolution de référence de 72 dpi nous donne un pixel de $1"/72$ (un pouce divisé par 72) soit $0.353mm$ [47][50].

II.2.3. L'image en niveau de gris :

Le niveau de gris est la valeur de l'intensité lumineuse en un point. La couleur du pixel peut prendre des valeurs allant du noir au blanc en passant par un nombre fini de niveaux intermédiaires.

Donc pour représenter les images en niveaux de gris, on peut attribuer à chaque pixel de l'image une valeur correspondant à la quantité de lumière renvoyée. Cette valeur peut être comprise par exemple entre 0 et 255. Chaque pixel n'est donc plus représenté par un bit, mais par un octet. Pour cela, il faut que le matériel utilisé pour afficher l'image soit capable de produire les différents niveaux de gris correspondant. Le nombre de niveaux de gris dépend du nombre de bits utilisés pour décrire la "couleur" de chaque pixel de l'image. Plus ce nombre est important, plus les niveaux possibles sont nombreux [47] [51].

II.2.4. L'image en couleurs :

Même s'il est parfois utile de pouvoir représenter des images en noir et blanc, les applications multimédias utilisent le plus souvent des images en couleurs. La représentation des couleurs s'effectue de la même manière que les images monochromes avec cependant quelques particularités [52]. En effet, il faut tout d'abord choisir un modèle de représentation. On peut représenter les couleurs à l'aide de leurs composantes primaires. Les systèmes émettant de la lumière sont basés sur le principe de la synthèse additive : les couleurs sont composées d'un mélange de rouge, vert et bleu (*RGB*) [47] [51].

Le modèle *RGB* (*RVB*) représentant toutes les couleurs par l'addition de trois composantes fondamentales, n'est pas le seul possible. Il en existe de nombreux autres. L'un d'eux est particulièrement important, il consiste à séparer les informations de couleurs (chrominance) et les informations d'intensité lumineuse (luminance). La chrominance est représentée par deux valeurs et la luminance par une valeur [47] [53].

La représentation en couleurs réelles consiste à utiliser 24 $bits$ pour chaque point de l'image. Huit bits sont employés pour décrire la composante rouge (*R*), huit pour le vert (*V*) et huit pour le bleu (*B*). Il est ainsi possible de représenter environ 16,7 $millions$ de couleurs différentes simultanément.

Cela est cependant théorique, car aucun écran n'est capable d'afficher 16 $millions$ de points. Dans la haute résolution ($1600\ x\ 1200$), l'écran n'affiche que 1 920 000 points. Par ailleurs, l'œil humain n'est pas capable de distinguer autant de couleurs [47] [54].

II.3. La compression en imagerie :

II.3.1. Elimination de l'information redondante :

Sur les images fixes, les méthodes de compression se basent sur la corrélation spatiale entre pixels (redondances spatiales). Pour la vidéo, on peut supprimer deux types de redondances :

- Les redondances spatiales, on utilisera pour se faire des techniques de transformées en fréquence de type *DCT* (transformée en cosinus discrète).

- les redondances temporelles (au moyen de techniques d'estimation et compensation de mouvements).

Les normes de compression d'images utilisées dans la visio-phonie et la visio-conférence *H.261* et *H.263* ainsi que *MPEG1* et *MPEG2* utilisent la *DCT* et la compensation de mouvements. On utilise également la redondance des valeurs à transmettre en attribuant aux valeurs les plus fréquentes les mots de code les plus courts, ce qui contribue à la réduction des informations à transmettre. Ces techniques de codes à longueurs variables sont réversibles et sont également utilisées pour la compression de fichiers informatiques.

Pour rentrer dans les détails, sachez que de nombreux compacteurs recourent à l'algorithme de compression Huffman qui code les données selon leur récurrence statistique :

Plus une séquence de données se répète dans le fichier, plus son code sera court. Á l'inverse, les séquences plus rares sont dotées d'un code plus long.

Á la décompression, tous les codes sont confrontés à une table de correspondance qui permet de réattribuer chaque séquence de données au code correspondant [47].

II.3.2. Compression sans distorsion :

II.3.2.1. Codage Shannon-Fano :

Shannon du laboratoire *Bells* et R. M. Fano du *MIT* ont développé à peu près en même temps une méthode de codage basée sur de simples connaissances de la probabilité d'occurrence de chaque symbole dans le message [55].

Le procédé de Shannon-Fano construit un arbre descendant à partir de la racine, par divisions successives. Le classement des fréquences se fait par ordre décroissant, ce qui suppose une première lecture du fichier et la sauvegarde de l'en-tête [47].

II.3.2.2. Codage Huffman :

Le codage de Huffman [56] crée des codes à longueurs variables sur un nombre entier de bits. L'algorithme considère chaque message à coder comme étant une feuille d'un arbre qui reste à construire. L'idée est d'attribuer aux deux messages de plus faibles probabilités, les mots codés les plus longs. Ces deux mots codés ne se différencient que par leur dernier bit.

Contrairement au codage de Shannon-Fano qui part de la racine des feuilles de l'arbre et, par fusions successives, remonte vers la racine. Le codage de Huffman consiste à coder les symboles par une représentation de bits à longueur variable. Les symboles ayant la probabilité d'apparition forte sont codés avec des chaines de bits plus courtes, tandis que les symboles dont la probabilité d'apparition est faible sont codés par des chaines plus longues. Le code d'un symbole ne doit pas être le préfixe d'un autre code. Cette propriété est admise afin que la reconnaissance soit possible. Pour représenter le codage de Huffman, on utilise l'arbre binaire [47].

II.3.2.3. Codage RLC (Run Length Coding) :

Plutôt que de coder seulement le message lui-même, il est plus intéressant de coder un message contenant une suite d'éléments répétitifs par "un couple répétition et valeur". Le codage *RLC* consiste en effet à coder un élément du message par sa valeur de répétition [47].

II.3.2.4. Codage arithmétique :

Contrairement aux algorithmes de Huffman et de Shannon-Fano qui associent à des symboles des motifs binaires dont la taille dépend de leur distribution.

Le codeur arithmétique traite le fichier dans son ensemble, en lui associant un unique nombre décimal rationnel. Ce nombre compris entre 0 et 1, possède d'autant moins de chiffres après la virgule que le fichier dont il est redondant. Ces chiffres décimaux dépendent non seulement des symboles du fichier dans l'ordre où ils apparaissent, mais aussi de leur distribution statistique [47].

II.3.3. Compression avec distorsion :

La figure 2.4 représente le schéma classique d'un système de compression et décompression. Dans un premier temps, afin de mieux compacter l'information, la source est transformée en groupe de coefficients. Les transformations les plus utilisées, que ce soit pour les images fixes ou les séquences d'images, sont la transformée en cosinus discrète *DCT,* la Transformée en ondelettes discrète *DWT* ou la décomposition Pyramidale.

Dans un second temps, les coefficients obtenus après la transformation sont quantifiés. La phase de quantification introduit l'erreur dans le système de codage. La dernière étape consiste à coder les coefficients quantifiés par le codage entropique [47].

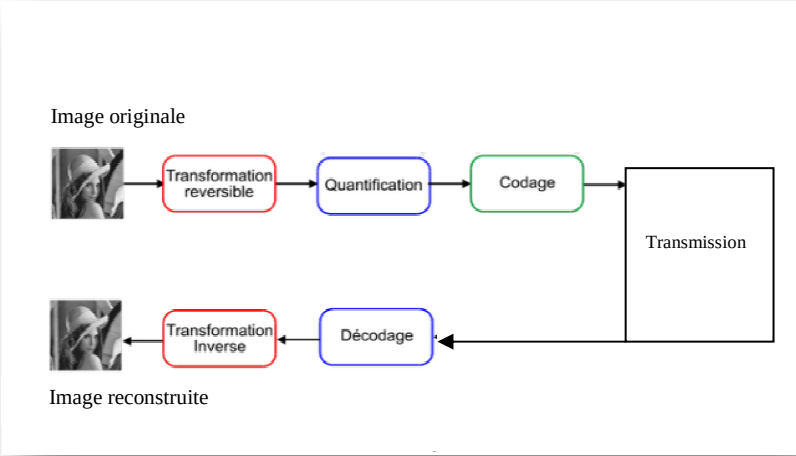


Figure 2.4. Schéma classique d'un système de compression et décompression.

II.3.3.1. Codage par quantification :

La quantification est l'une des sources de perte d'information dans le système de compression. Son rôle est en effet de réduire le nombre de bits nécessaire à la représentation de l'information. Elle est réalisée avec la prise en compte de l'aspect visuel, ce qui permet de déterminer la distorsion tolérable à apporter au signal à coder.

- ***Codage par prédiction***

C'est la technique de compression la plus ancienne. On prédit la valeur du pixel à partir de la valeur précédemment codée. La prédiction peut se faire au moyen 12 de l'histogramme de l'image. Seul l'écart entre la valeur réelle et la valeur prédite est quantifié puis codé et envoyé au décodeur. On peut réaliser la prédiction, au sein de l'image elle-même ainsi qu'entre images d'une séquence. Cette dernière est connue sous le nom de prédiction par compensation de mouvement. Le codage par prédiction est utilisé dans le codage *DPCM* (Differential Pulse Code Modulation).

- ***Quantification scalaire***

La quantification scalaire est réalisée indépendamment pour chaque élément. D'une manière générale, on peut la définir comme étant l'association de chaque valeur réelle x, à une autre valeur q qui appartient à un ensemble fini de valeurs. La valeur q peut être exprimée en fonction de la troncature utilisée : soit par l'arrondi supérieur, l'arrondi inférieur, ou l'arrondi le plus proche. On appelle le "pas de quantification Δ" l'écart entre chaque valeur q. Arrondir la valeur x provoque une erreur de quantification, appelé le "bruit de quantification". La valeur classique de ce dernier est $\Delta^2/12$. La procédure suivante définit la réalisation d'une quantification scalaire. Soit X l'ensemble d'éléments d'entrée de taille N:

1. Echantillonner X en sous-intervalles $\{[x_n, x_{n+1}[/n \in \{0 \dots N-1\}\}$.

2. Associer à chaque intervalle $[x_n, x_{n+1}[$ une valeur q.

3. Coder une donnée $x \in X$ par q si $x \in [x_n, x_{n+1}[$.

Si Δ est constant, on parle d'une quantification uniforme. Sinon elle est dite non-uniforme.

- ***Quantification vectorielle***

La quantification s'effectue sur un groupe d'éléments de la source, représenté par un vecteur $\vec{x}$ de dimension n. La source est constituée par un ensemble fini de vecteurs $\vec{x}$. La QV (Quantification Vectorielle) consiste alors à remplacer le vecteur $\vec{x}$ par un vecteur $\vec{y}$ de même dimension appartenant à un dictionnaire [57]. Le dictionnaire est un ensemble fini de vecteurs codes. Un vecteur $\vec{x}$ codé, appelé classe est obtenu en faisant la moyenne itérative de vecteurs $\vec{y}$. La quantification vectorielle se décompose en général en deux parties : le processus de codage et le processus de décodage. Le processus de codage cherche l'adresse du vecteur $\vec{y}$ correspondant dans le dictionnaire et l'envoie au récepteur. Le décodeur, quant à lui, dispose d'une réplique du dictionnaire et consulte celle-ci afin de reconstruire le vecteur code correspondant à l'adresse reçue. Dans le domaine de la compression d'images fixes, des travaux ont été proposées pour élaborer le dictionnaire [58][59].

II.3.3.2. Codage par transformée :

La transformation des données d'entrée est faite afin de mieux compacter l'énergie de la transformée d'image sur un nombre faible de coefficients. La transformation a pour objet de décorréler les pixels d'image. On opère la transformation sur un bloc unitaire ou directement sur l'image entière.

- ***Transformée de Fourier Discrète TFD***

La *TFD* permet de passer du domaine spatial au domaine fréquentiel. La *TFD* d'un signal discret x_n s'exprime par :

$$X_k = \sum_{n=0}^{N-1} x_n exp\left(-i\frac{2\pi nk}{N}\right) \quad (2.1)$$

La fonction inverse permettant de remonter au signal original $x_n$connaissant sa transformée X_k est :

$$x_n = \sum_{n=0}^{N-1} X_k exp\left(+i\frac{2\pi nk}{N}\right) \quad (2.2)$$

- ***Transformée en Cosinus Discrète DCT***

La *DCT* est une variante de la *TFD* et il en existe plusieurs. Dans le cadre du signal monodimensionnel, la plus connue est celle-ci :

$$X_k = \sum_{n=0}^{N-1} x_n cos\left(\frac{\pi}{N}\left(n+\frac{1}{2}\right)k\right) \quad (2.3)$$

La *DCT* est très utilisée dans la compression d'image fixe, notamment dans la norme *JPEG*. Dans cette application la *DCT* s'effectue sur un bloc de pixels 8×8 et qui s'exprime par :

$$X_{u,v} = \frac{1}{4}C(u)C(v)\sum_{n=0}^{7}\sum_{n=0}^{7} cos\left(\frac{(2n+1)u\pi}{16}\right) cos\left(\frac{(2m+1)v\pi}{16}\right) \quad (2.4)$$

avec

$$u, v = \{0 \dots 7\}\ et\ \ C(w) = \begin{cases} \frac{1}{\sqrt{2}} & si\ w = 0 \\ 1 & sinon \end{cases}$$

- ***Transformée en Ondelettes Discrète DWT***

Contrairement à la transformée de Fourier, la transformée en ondelettes permet de déterminer les différentes composantes fréquentielles d'un signal donné, ainsi que leur localisation spatiale ou temporelle. Par définition, les ondelettes sont des fonctions gérées à partir d'une ondelette mère ψ, par dilatations et translations. Ainsi, la décomposition en ondelettes fait intervenir deux paramètres qui sont le facteur d'échelle s et le facteur de translation τ [60].

Le paramètre d'échelle s permet d'obtenir des ondelettes à partir d'une ondelette mère, des ondelettes comprimées ainsi que des ondelettes dilatées. Les ondelettes comprimées sont utilisées pour déterminer les composantes hautes fréquences tandis que les ondelettes dilatées permettent de déterminer les composantes basses fréquences.

D'après Mallat [61], dans le cas d'un espace à deux dimensions, la fonction d'ondelette monodimensionnelle ψ avec la fonction d'échelle peut être séparée en trois fonctions d'ondelettes distinctes (ψ_1, ψ_2, ψ_3) :

– $\psi_1(x,y) = \varphi(x)\psi(y)$ exprime les variations selon l'axe horizontal.
– $\psi_2(x,y) = \psi(x)\varphi(y)$ exprime les variations selon l'axe vertical.
– $\psi_3(x,y) = \psi(x)\psi(y)$ exprime les variations selon les deux axes (diagonal).

Il existe une technique appelée "transformée standard", qui permet de réaliser la transformée en ondelettes à deux dimensions. En effet la technique consiste à calculer la transformée en ondelettes mono-dimensionnelle pour chaque élément de la ligne puis d'appliquer la transformée en ondelettes mono-dimensionnelle pour chaque élément de la colonne. La transformation en ondelettes peut être considérée comme une projection du signal f sur ces différents sous-espaces. Cette décomposition peut s'obtenir simplement par des convolutions du signal avec des filtres miroirs en quadrature. Á chaque niveau, on filtre le signal par des filtres d'analyse passe-haut h_H et passe-bas h_L, et on sous-échantillonne avec un facteur de 2. Ainsi, après la décomposition, on obtient quatre matrices chacune de taille quatre fois plus petite. D'où l'obtention de la

représentation multi-résolution dans laquelle on a quatre sous-bandes. On les désigne par LL, LH, HL et HH. La lettre H correspond au filtrage passe-haut et la lettre L à celui du passe-bas appliqué de façon séparable sur les lignes et les colonnes.

Ainsi, pour reconstruire le signal original, on applique une succession de filtres conjugués, filtres de synthèse, en sur-échantillonnant avec un facteur de 2 le signal décomposé. Le filtre d'analyse est représenté par le couple (H_0, H_1) tandis que le couple (G_0, G_1) représente le couple du filtre de synthèse. Afin d'obtenir une reconstruction parfaite, les deux bancs de filtres d'analyse et de synthèse doivent satisfaire la relation ci-après. Soit $H_L(z)$, $G_L(z)$,$H_H(z)$, et $G_H(z)$, les transformées en Z, des filtres H_0,G_0, H_1, G_1 respectivement :

$$H_L(z)G_L(z) + H_H(z)G_H(z) = 2 \tag{2.5}$$

$$H_L(-z)G_L(z) + H_H(-z)G_H(z) = 0 \tag{2.6}$$

II.3.4. Normes de compression de l'image fixe :

De nombreux organismes internationaux s'occupent de la normalisation des systèmes de codage pour des images fixes :

l'Organisation Internationale de Normalisation (*ISO*) et la Commission Electrotechnique Internationale (*IEC*). On s'intéressera aux travaux du groupe *JPEG* qui est chargé des spécifications de l'évolution du standard d'image fixe de *l'ISO/IEC*.

Les normes de cette famille sont connues sous les noms *JPEG* et *JPEG*2000. La technique utilisée pour l'encodage d'images fixes peut être ramenée à deux grandes familles : méthode par transformée et technique structurelle. L'approche structurelle emploie les techniques qui manipulent la valeur d'intensité du pixel dans l'image. Les deux grandes familles peuvent être utilisées ensemble pour un système de codage. Autrement dit, il n'existe pas de frontière entre l'approche par transformée et l'approche structurelle. On appelle taux de compression le rapport entre l'image brute et l'image compressée.

II.3.4.1. La norme JPEG :

Norme de compression d'image fixe, établie en 1991, *JPEG* permet la compression sans perte et avec perte d'information. C'est une norme qui a eu beaucoup de succès depuis plusieurs années et que nous utilisons toujours aujourd'hui.

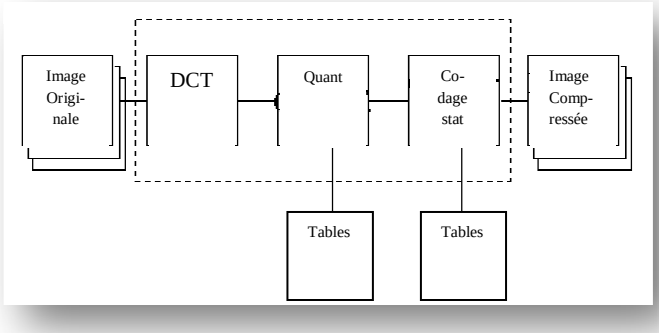


Figure 2.5. Principe de l'algorithme *JPEG* avec perte.

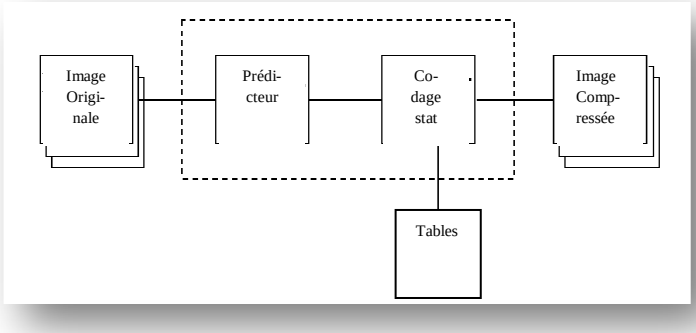


Figure 2.6. Principe de l'algorithme *JPEG* sans perte.

L'encodage par *JPEG* est réalisé en quatre étapes successivement : transformée en cosinus discrète *DCT*, phase de quantification, réorganisation en zig-zag, suivie de codage *RLC* et codage de Huffman.

- ***DCT*** La première étape consiste à découper l'image originale en blocs de taille 8×8. Ensuite pour chaque bloc, on applique la *DCT*. On a donc une matrice dont les composantes sont les coefficients de la transformée *DCT*. Ainsi, on obtient deux sortes de coefficients, *DC* et *AC*. Le coefficient *DC* représente la moyenne des pixels appartenant au bloc courant, les éléments restants sont des coefficients *AC*.

- ***Quantification*** C'est la deuxième étape. Cette phase consiste à donner les valeurs approximatives de la matrice précédente. Pour cela, une matrice de quantification est construite.

- ***Réorganisation en zig-zag*** L'objet de cette troisième étape est d'obtenir le maximum de suites de zéros. En effet, la norme *JPEG* traite les valeurs zéro de la matrice quantifiée en raison de leur nombre important. Ceci concerne seulement les coefficients *AC* du bloc. Le coefficient *DC*, quant à lui, est codé par le codage différentiel. Ce dernier consiste à coder la différence entre les deux coefficients *DC* successifs courant et précédent.

- ***Codage entropique*** La dernière étape est le codage de Huffman des coefficients obtenus précédemment. Ainsi, on a une image compressée en *JPEG*.

- ***Limite du standard JPEG*** La norme *JPEG* ne permet pas un taux de compression élevé supérieur à 32. Au-delà, des artefacts très gênants sont brusquement présents dans l'image codée.

*II.3.4.2. La norme JPEG*2000 *:*

*JPEG*2000 est une norme de compression d'image fixe. Introduite en mars 1997 par l'*ISO/IEC*, elle est devenue standard en décembre 2000. L'objectif de la norme est de compléter la performance du standard *JPEG* mais sans la remplacer [62].

Par rapport à *JPEG*, à la place de la transformée *DCT*, le standard *JPEG*2000 utilise la transformée en ondelettes discrète *DWT*. Au lieu du codage de Huffman, la norme *JPEG*2000 emploie un codage entropique basé sur le codage arithmétique et la *DWT* permet la scalabilité, obtenue grâce à la représentation multirésolution.

L'image originale est découpée en petits blocs rectangulaires appelés tuiles dont les dimensions correspondent à des puissances de 2 $(16, 32, 64, etc)$. Ainsi obtenues, les tuiles ont les mêmes dimensions, mises à part, éventuellement, les tuiles constituant les bords à droite et en bas de l'image. Chaque tuile est codée indépendamment, ce qui réduit la quantité de mémoire nécessaire à l'encodage. L'étape suivante est la transformée couleur. Cette phase permet de décorréler les composantes couleurs. Le standard utilise deux types d'ondelettes : les ondelettes de Daubechies (9,7) et de Daubechies (5,3).

Ces deux ondelettes sont choisies selon le type de la compression souhaitée, sans perte ou avec perte. Les ondelettes de Daubechies (5,3) sont utilisées pour la compression sans ou avec perte, tandis que les ondelettes de Daubechies (9,7) sont utilisées uniquement pour la compression avec perte.

- ✓ Dans le système de compression en général, la quantification est l'étape qui réalise la compression des données. Elle permet en effet la troncature de tous les coefficients d'ondelettes dans chaque sous-bande.

II.4. La vidéo Numérique :

Une vidéo est une succession d'images à une certaine cadence. L'œil humain a comme caractéristique d'être capable de distinguer environ 20 images par seconde. Ainsi, en affichant plus de 20 images par seconde, il est possible de tromper l'œil et de lui faire croire à une image animée. La vidéo numérique consiste à afficher une succession d'images numériques. D'autre part la vidéo au sens multimédia du terme est généralement accompagnée de son, c'est-à-dire de données audio.

Puisqu'il s'agit d'images numériques affichées à une certaine cadence, il est possible de connaître le débit nécessaire pour l'affichage d'une vidéo. Ainsi le débit nécessaire pour afficher une vidéo est égal à la taille d'une image que multiplie le nombre d'images par seconde. Soit une image couleur $(24\ bits)$ ayant une définition de 640 pixels par 480. Pour afficher correctement une vidéo possédant cette définition il est nécessaire d'afficher au moins 30 images par seconde, c'est-à dire un débit égal à $(921ko \times 30) = 27Mo/s$ [63].

II.5. La Norme de compression vidéo MPEG :

Les efforts développés par les équipes du *CCITT* pour le *H.261* ont été utilisés comme point de départ pour le développement d'un standard de codage d'images animées par l'*ISO*, ce standard s'intitule *MPEG* pour "Moving Pictures Experts Group" [64]. Contrairement au *H.261* en première phase, *MPEG* est destinée au codage des images animées en vue de leur stockage sur les supports numériques d'où les contraintes plus souples que celles du *H.261* [65].

La première phase de *MPEG* intitulée *MPEG-1* spécifie une compression du signal vidéo à un débit de 1.5 $Mbits/s$. Les deux autres phases ont pour but d'améliorer la qualité du codage vidéo en sacrifiant une augmentation de débit, *MPEG-2* est destinée à la compression du signal vidéo à des débits d'ordre de 10 $Mbits/s$, *MPEG-3* étant destinée à la télévision haute définition à des débits de 30 à 40 $Mbits/s$. Une quatrième phase, *MPEG-4* est destinée au codage d'images animées à très faibles débits (10 $Mbits/s$).

Le *MPEG-7*, une phase visant à fournir une représentation standard des données audio et visuelles afin de rendre possible la recherche d'information dans de tels flux de données. Ce standard est ainsi également intitulé Multimédia Content Description Interface (MCDI).

II.6. Le codec H.264 :

La norme de codage vidéo *H.264/ AVC* [66] s'inscrit dans la continuité des normes précédentes de la famille *MPEG*. Cette norme a été mise en place en 2003 par le *JVT*, l'équipe de travail conjointe de l'*ITU-T* (*CCITT*) et de l'organisation internationale de normalisation *ISO*. L'apport principal de ce système de codage est son efficacité par rapport à ces prédécesseurs, permettant un gain en débit de 50% par rapport à *MPEG-2* à qualité constante. La norme *H.264/AVC* a les mêmes éléments fonctionnels de base que les autres standards tels que la transformation pour la réduction de la corrélation spatiale, la quantification pour le contrôle de débit binaire, la prédiction compensée en mouvement pour diminuer la corrélation temporelle et le codage entropique pour réduire la corrélation statistique.
Cependant, afin d'accomplir une meilleure performance de codage, les changements importants dans *H.264/AVC* se produisent dans les détails de

chaque élément fonctionnel. Améliorer l'efficacité du codage vient aux dépens de la complexité calculatoire au sein du codeur/décodeur. *H.264/AVC* adopte des méthodes pour réduire la complexité d'implantation. Les conditions des canaux à bruit tels que les réseaux sans fil, obstruent la réception parfaite du train binaire par le décodeur. Un décodage incorrect dégrade la qualité subjective de l'image. Pour des applications particulières telles que la vidéo-conférence et la vidéo-téléphonie, le codage haute qualité pour la transmission et le vidéo streaming sur les réseaux par paquets est nécessaire.

II.6.1. L'encodeur H.264 :

II.6.1.1. Principe de fonctionnement :

Le fonctionnement d'un codeur *H.264/AVC* est illustré sur figure 2.7. Le codage d'une image commence par son découpage en blocs de taille 16×16 appelés macroblocs.

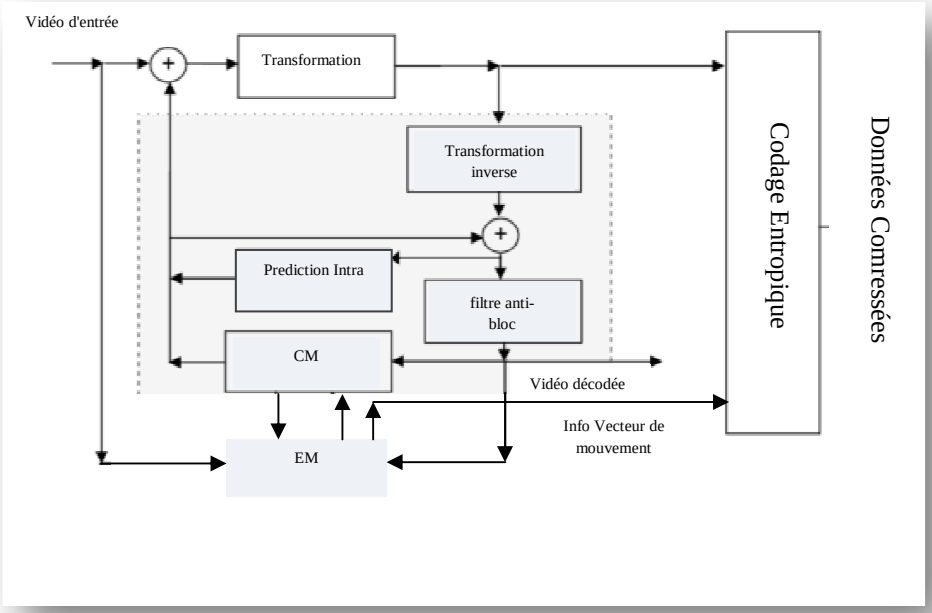


Figure 2.7. Schéma fonctionnel du codage *H.264/AVC*.

Selon le type de l'image, ces derniers sont codés en mode intra ou inter. Le codage intra n'utilise que des échantillons de blocs déjà codés dans la même image. Le codage inter utilise des images décodées (images de référence) pour obtenir la meilleure prédiction d'un bloc de l'image codée. Ensuite, la différence entre le bloc prédit (intra ou inter) et le bloc original est calculée et une transformée entière est appliquée à cette information résiduelle. Les coefficients de la transformée sont alors quantifiés et codés par un algorithme de codage entropique.

Chacun des macroblocs subit alors les étapes de prédiction générant un résiduel de pixels à partir des données causales, transformation et quantification du résiduel puis finalement encodage à l'aide d'un codeur entropique des résiduels de pixels et des informations de signalisation nécessaires au décodage du flux. Le flux binaire obtenu en sortie de l'encodeur peut alors être lu par un décodeur qui restitue une vidéo en répétant ces étapes de façon inverse et en incluant une étape de filtre anti-bloc (en anglais, deblocking filter) permettant de diminuer les dégradations du signal. Nous allons détailler dans les sections qui suivent, les modules du codeur *H.264*.

II.6.1.2. Prédiction H.264 :

H.264/AVC supporte plusieurs types de prédiction : prédiction intra utilisant des données dans l'image actuelle, inter prédiction en utilisant une prédiction en compensation de mouvement à partir d'images précédemment codées, prédiction de bloc de différentes tailles, de multiples images de référence et des modes spéciaux tels que la prédiction directe.

Ces caractéristiques donnent à l'encodeur *H.264* une grande flexibilité dans le processus de prédiction.

En sélectionnant les meilleures types de prédiction pour un macro-bloc individuel, un codeur peut réduire la taille du résiduel pour produire un flux binaire fortement comprimé.

Nous commençons par un aperçu de la prédiction de macro-bloc, suivie d'un aperçu détaillé de prédiction intra et inter en *H.264 / AVC* [135].

A. Prédiction de Macro-blocs :

- ✓ Un macro-bloc de type **I** (Intra) est prédit utilisant une prédiction intra à partir des échantillons voisins dans l'image actuelle.

- ✓ Un macro-bloc **P** est prédit à partir des échantillons dans une trame précédemment codée qui peut être avant ou après l'image actuelle dans l'ordre d'affichage. des sections rectangulaires différentes dans un macro-bloc **P** peuvent être prédites à partir de différentes images de référence.

- ✓ Chaque partition dans un macro-bloc **B** est prédite à partir des échantillons dans une ou deux images précédemment codées, par exemple, un "passé" et un "futur".

B. Prédiction Intra :

Un macro-bloc **I** est codé sans se référer à des données en dehors de la tranche en cours. un macro-bloc **I** peut se produire dans tout type de tranche. Chaque macro-bloc dans une tranche I est un macro-bloc **I**. Ces macro-blocs sont codés utilisant une prédiction intra, i.e. la prédiction à partir des données précédemment codées dans la même tranche. Pour un bloc typique d'échantillons de la luminance ou la chrominance (luma ou chroma), il y a une corrélation relativement élevée entre les échantillons dans le bloc et les échantillons qui sont immédiatement adjacents au bloc. La prédiction intra utilise donc les échantillons provenant, des blocs adjacents précédemment codés pour prédire les valeurs dans le bloc courant.

Dans un macro-bloc intra, il y a trois choix de taille de bloc pour la prédiction intra pour la composante de la luminance, à savoir 16×16, 8×8 ou 4×4. Un seul bloc de prédiction est généré pour chaque composante de chrominance. Le tableau 2.1 explique les différents types de la prédiction Intra pour la luminance et la chrominance.

Modes de prédiction Intra	Notes
Intra 16×16 (luma)	Un bloc de prédiction P 16×16 est généré. 4 modes de prédiction possible.
Intra 8×8 (luma)	Un bloc de prédiction P 8×8 est généré pour chaque bloc 8×8 luma. 9 modes de prédiction possible.
Intra 4×4 (luma)	Un bloc de prédiction P 4×4 est généré pour chaque bloc 4×4 luma. 9 modes de prédiction possible
Chroma	Un bloc de prédiction P est généré pour chaque composante de chrominance. 4 modes de prédiction possible.

Tableau 2.1. Types de prédiction Intra [135].

Lorsque la taille du bloc 4×4 ou 8×8 est choisi pour la composante de luminance, l'intra prédiction est créée à partir des échantillons directement au-dessus du bloc courant, directement à gauche du bloc, ci-dessus et à gauche et ci-dessus et à droite. Pour un bloc de luminance de taille $16 \, x \, 16$ ou un bloc de chrominance, la prédiction est créée à partir des échantillons directement vers la gauche ou au-dessus du bloc actuel, ou d'une combinaison de ceux-ci.

Le codeur sélectionne un mode intra pour le bloc courant à partir des modes de prédiction disponibles. Le choix du mode intra est communiqué au décodeur. Chaque composante de chrominance d'un macro-bloc est prédite à partir des échantillons de chrominance codés précédemment ci-dessus et/ou vers la gauche.

H.264/AVC offre un ensemble riche de modèles de prédiction Intra (9 modes de prédiction Intra 4×4, 4 modes de prédiction Intra 16×16 et 4 modes de prédiction des blocs de chrominance). Avec l'emploi de ces techniques avancées de prédiction Intra, l'efficacité du codage a été significativement améliorée.

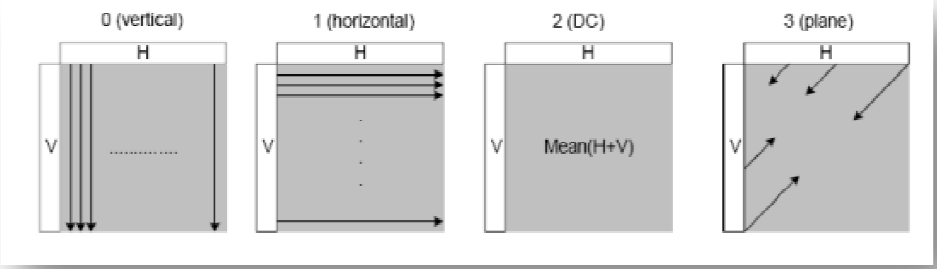


Figure 2.8. Modes d'intra-prédiction d'un bloc de luminance 16 × 16 [135].

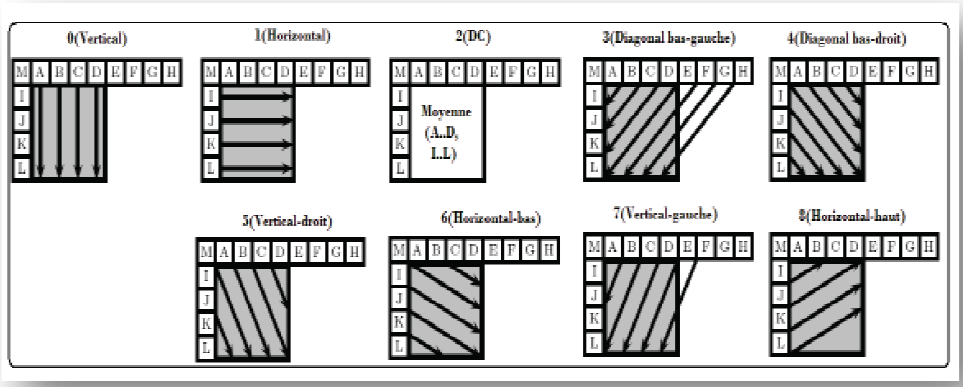


Figure 2.9. Prédiction d'un bloc 4 × 4 en mode intra [135].

Les quatre modes de prédiction de chrominance sont très similaires aux mode de prédiction 16 × 16 décrits et illustré précédemment, sauf que la numérotation des modes est différente. Les modes sont *DC* (mode 0), horizontal (mode 1), vertical (mode 2) et le plane (mode 3).

C. Prédiction Inter :

La prédiction Inter est le processus de prédiction d'un bloc d'échantillons de luminance et chrominance d'une image qui a déjà été codée et transmise (une image de référence).

Cela consiste à sélectionner une région de prédiction, qui génère un bloc de prédiction et en soustrayant ceci à partir du bloc original afin de former un résiduel qui est ensuite codé et transmis. L'image de référence est choisie parmi une liste d'images précédemment codées, stockées dans une mémoire tampon des images décodées, ce qui peut inclure des images avant et après l'image actuelle dans l'ordre d'affichage. Le décalage entre la position actuelle de la partition et de la région de prédiction dans l'image de référence est un vecteur de mouvement.

La prédiction Inter utilise un ensemble de mécanismes tels que : la compensation de mouvement, la prédiction des vecteurs de mouvement. Le principe de l'estimation de mouvement est de capturer les mouvements des *MB* dans les images au cours du temps, elle peut être faite de deux façons : "avant" ou "arrière" [67]. La compensation de mouvement est l'opération permettant de construire l'image courante macro-bloc par macro-bloc à partir d'une image de référence et d'information du vecteur de mouvement, elle permet cette reconstruction en déduisant le macro-bloc prédit dans l'image de référence puis en le soustrayant du macro-bloc de l'image courante pour obtenir le macro-bloc résiduel.

Les tailles de blocs disponibles sont : 16×16, 16×8, 8×16 et 8×8 pouvant lui-même être partitionné en8×4, 4×8 et 4×4. Le standard utilise de plus une représentation sous-pixellique du mouvement qui peut ainsi être réalisée jusqu'au $1/4$ de pixel. Un mode particulier est défini en Inter, il s'agit du mode Skip qui s'applique sur les blocs de taille 16×16 et a la particularité de ne nécessiter aucune information additionnelle (ni vecteur de mouvement, ni résiduels) [135].

II.6.1.3. Transformation et quantification :

Après l'étape de prédiction, les données organisées sous forme de macro-blocs sont transformées et quantifiées, et cela avant l'opération finale de codage et de réarrangement en bitstream. De même dans un décodeur, une transformée inverse est préalable à la reconstruction et l'affichage. Plusieurs transformations sont spécifiées dans la norme *H.264*. En effet, dans un codec *H.264*, des transformations et des quantifications directes et inverses sont appliquées sur des blocs 4×4 pixels de luminance et de chrominance.

Les sections suivantes décrivent les processus de transformation et de quantification direct et inverse pour les coefficients de chrominance et de luminance.

Le processus de transformée et quantification directe ne sont pas standardisés, mais des procédés équivalents peuvent être dérivés de la transformation et la quantification inverse (figure 2.11). Dans un encodeur *H.264*, un bloc de coefficients résiduels est transformée et quantifiée (figure 2.10). Le Cœur de la transformée, est une transformée 4×4 ou 8×8 entière. Dans certains cas, une partie de la sortie de cette transformation entière est encore transformée utilisant une transformée de Hadamard, puis les coefficients transformés sont quantifiés. Les procédés inverses correspondants sont présentés sur la figure 2.11.

La transformée *DC*, si elle est présente, est effectuée avant la remise à l'échelle. Les coefficients remises à l'échelle sont appliqués aux transformées entières inverses 4×4 ou 8×8.

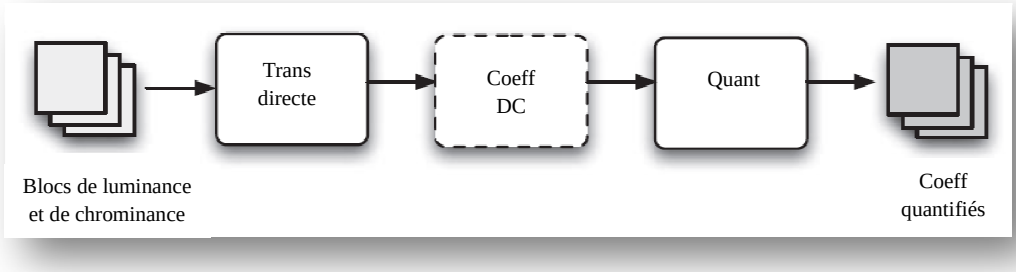


Figure 2.10. Transformée directe et quantification.

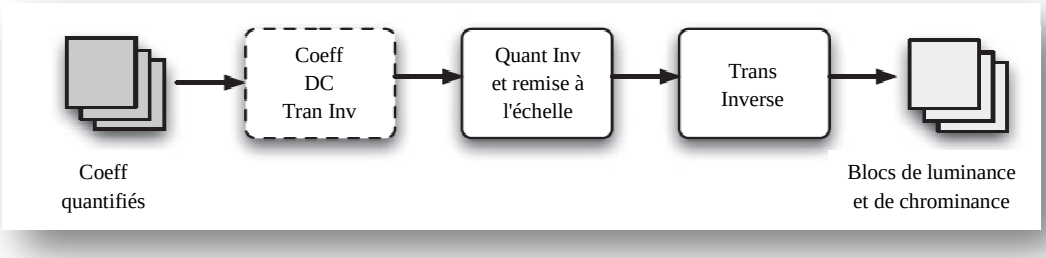


Figure 2.11. Transformée inverse et remise à l'échelle.

II.6.1.4. Codage entropique :

Dans la norme *H.264* le codage entropique peut être réalisé de trois manières différentes, une première méthode utilise une table universelle de mots de code *UVLC* (Unified Variable Lenght Coding) la deuxième méthode s'agit d'une part d'un codage à longueur variable *CAVLC* (Contest Adaptative Variable Lenght Coding) et un codage arithmétique adaptatif contextuel *CABAC* (Context Adaptative Binary Arithmetic Coding) [68].

- ***Codage du Golomb exponentiels (UVLC) :***

Les codes du Golomb exponentiels ($Exp-Golomb$) sont des codes de longueur variable avec une construction régulière. Il est facile de l'examen des premiers mots de code dans le tableau 2.2 de déduire qu'ils sont de la forme :

$$[M\ z\acute{e}ros][1][INFO] \tag{2.7}$$

$INFO$ est un champ binaire d'informations, les mots de codes 1 et 2 ont un seul champ de bit $INFO$, les mots de codes 3 et 6 ont deux champs de bit $INFO$. La longueur du code est $(2M+1)$ bits et chaque code peut être construit par le codeur en se basant sur l'indice $code\_num$:

$$M = floor(log_2[code\_num + 1]) \tag{2.8}$$

$$INFO = code\_num + 1 - 2M \tag{2.9}$$

$code\_num$	$Mot\ de\ code$
0	1
1	010
2	011
3	00100
4	00101
5	00110
6	00111
7	0001000
8	0001001
...	...

Tableau 2.2. Code Exponential Golomb [135].

Un mot de code peut être décodé comme suit :

1. Lire les premiers M zéros précédents le 1.

2. Lire les M bits du champ $INFO$.

3.$Code\_num = 2M + INFO - 1$.

- ***Codage CAVLC :***

L'algorithme *CAVLC* est utilisé pour coder des blocs résiduels de luminance et de chrominance 4×4 ou 2×2 transformés et quantifiés parcourus en zig-zag. *CAVLC* est conçu pour tirer avantage des caractéristiques des blocs quantifiés :

1. Les blocs de coefficients *DCT* après quantification sont composés en grande partie de zéros, *CAVLC* utilise une syntaxe afin de réduire au mieux de longue séquence de zéros.

2. Après le parcours en zig-zag, les coefficients non nuls situés à la fin de la chaine sont souvent des ± 1, *CAVLC* permet de coder efficacement ces suites de ± 1.

3. Le nombre de coefficients non nuls dans les blocs voisins est corrélé, de ce fait *CAVLC* utilise des tables en fonction du nombre de coefficients non nuls des blocs voisins.

4. L'amplitude des coefficients non nuls tend à être plus faible vers les hautes fréquences.

- **Codage *CABAC* :**

Le *CABAC* est un codeur arithmétique binaire adaptatif utilisé dans le standard de codage vidéo *H.264*. En contrepartie, ce code est très sensible au bruit de transmission : une seule erreur peut conduire à une désynchronisation totale du codeur. Ce code permet d'allier une bonne efficacité en compression. Le *CABAC* détermine un contexte d'après la valeur à encoder et les valeurs précédemment encodées et met à jour des probabilités d'encoder un 0 ou un 1.

Ce codage arithmétique a permis de réduire encore le débit binaire de 10 à 15% des signaux de télévision d'une même qualité en le comparant au *CAVLC*.

II.6.2. Décodeur H.264 :

- ✓ Un décodeur vidéo *H.264* reçoit le train binaire compressé, décode chacun des éléments de syntaxe et extrait les informations à savoir, les coefficients de transformés et quantifiés et les informations de prédiction, etc. Cette information est ensuite utilisée pour le processus de décodage.

- ✓ Les coefficients quantifiés sont remises à l'échelle. Chaque coefficient est multiplié par une valeur entière *QP* pour rétablir son échelle d'origine.

- ✓ Ces coefficients remises à l'échelle sont combinés à une transformée inverse pour regénérer chaque bloc de données résiduel.

- ✓ Pour chaque macro-bloc, le décodeur forme une prédiction identique à celle créée par le codeur en utilisant une prédiction inter à partir des trames précédemment décodés ou une prédiction intra à partir des échantillons décodés dans l'image actuelle [135].

II.7. Conclusion

Ce deuxième chapitre a caractérisé généralement le codage d'image et vidéo avant de présenter plus en détails sur l'ensemble de concepts généraux de compression. Nous avons également mis en avant un aperçu sur l'image et la vidéo numérique. Nous avons ensuite présenté le codeur *H.264* et notamment avec les concepts utilisés en prédiction intra-images et inter-images. Plusieurs versions du codeur *H.264/AVC* ont vu le jour. Ces version se basent sur quatre principaux modules afin de réaliser la compression des données de la vidéo d'entrée, la prédiction temporelle est réalisée par une estimation et une compensation en mouvement de l'image de référence. Une transformation fréquentielle qui permet de décorréler spatialement les informations, qui est une approximation de la *DCT* calculée matriciellement en nombres entiers sur des blocs de taille 4×4 et/ou 8×8.

Une quantification, qui consiste à réduire l'amplitude des coefficients issus de la transformation *DCT*. Enfin un codage entropique est appliqué. Après cette étude détaillée de la norme de compression vidéo *H.264/AVC*, néanmoins. Nous aborderons dans le chapitre suivant la description des représentations parcimonieuses qui permettent de représenter un signal d'une manière plus sparse.

III.1. Codage à descriptions multiples :

III.1.1. Introduction :

Le codage à descriptions multiples a émergé comme un schéma attractif pour une transmission robuste sur les canaux non fiables. Il peut efficacement lutter contre la perte de paquets sans aucune retransmission répondant ainsi à la demande des services en temps réel. Le *MDC* encode le message source en plusieurs flux binaires portant des informations différentes qui peuvent ensuite être transmises sur des canaux différents. Si un seul canal fonctionne, les descriptions peuvent être décodées individuellement afin de garantir une fidélité minimale suffisante à la reconstruction. Toutefois, lorsque plusieurs canaux fonctionnent, les descriptions des canaux peuvent être combinées pour obtenir une reconstruction de plus grande fidélité. Au cours des dernières années, le *MDC* d'image fixe a attiré beaucoup d'attention, et plusieurs algorithmes ont été proposés [69][70][71].

Récemment, la version *MDC* basée sur l'échantillonnage est une technique populaire pour la conception de codage vidéo à descriptions multiples. Dans [72] un schéma *MDC* avec un pré et post traitement est présenté. La redondance est introduite par l'insertion des zéros dans le domaine *DCT* de chaque image et les descriptions multiples sont générées par sous-échantillonnage des images sur-échantillonnées. Par le sur-échantillonnage des images, seulement la corrélation Inta est augmentée pour améliorer la distorsion latérale, par contre la corrélation entre les images (corrélation en Inter) est complètement négligée.

Une autre approche *MDC* [73] consiste d'abord à sur-échantillonner l'image d'entrée pour augmenter la corrélation entre les échantillons adjacents, puis ils appliquent une décomposition sur l'image sur-échantillonnée pour générer des sous images qui peuvent être encodées et transmises séparément. Le décodeur décode simplement les paquets qu'il reçoit. Si toutes les descriptions sont reçues ou une description ou plus sont perdues, c'est l'étape de post-traitement qui combine les images reçues pour estimer les données perdues.

Dans un autre travail [74] Gallant et al applique cet approche à travers un pré et post-traitement des séquences vidéo sans aucune modification du codeur source ou de canal.

Ils utilisent le sur-échantillonnage pour rajouter de la redondance aux données de la vidéo suivi d'une décomposition en sous-images de tailles égales [136].

III.1.2. Présentation du problème :

En 1979, lors d'un atelier sur la théorie de Shannon, la question suivante a été posée par Gersho, Ozarow, Witsenhausen, Wolf, Wyner et Ziv [75] : Si une source d'information est décrite par deux descriptions distinctes, quelles sont les limitations en qualité de ces descriptions prises séparément et ensemble ? Les premiers résultats théoriques sont apparus dans les années 1980 [75][76]. Le problème de codage *MD* sera d'abord décrit avec deux descriptions illustré sur la figure 3.1. La source $\{X_k\}_{k=1}^{N}$est transmise sur deux canaux distincts. Les canaux ont deux états de fonctionnement possibles : soit ils sont en état de marche et transmettent sans erreur tous les symboles, soit ils sont défaillants et ne transmettent aucun symbole. Le récepteur connait l'état des canaux mais pas le transmetteur qui envoie systématiquement deux descriptions sur les deux canaux.

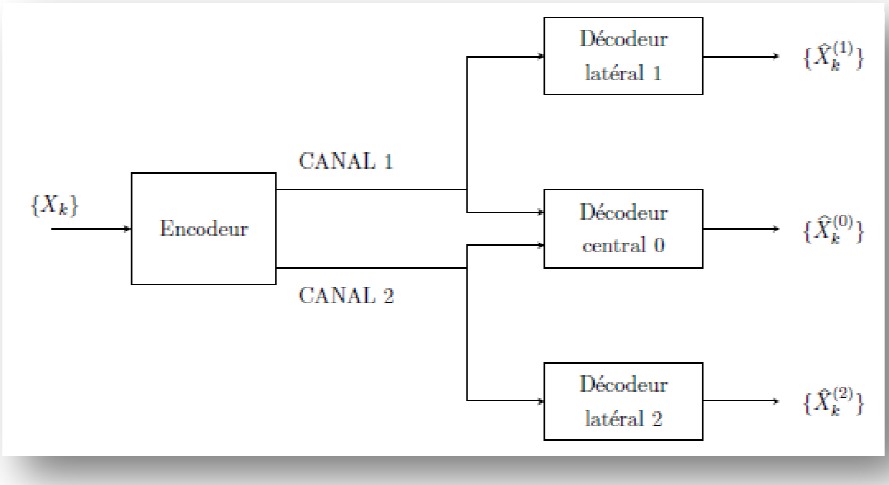


Figure 3.1. Schéma générique de codage *MD* avec deux descriptions [136].

Au récepteur, les trois décodeurs correspondent à trois situations différentes. Le décodeur central est utilisé quand les deux canaux sont en état de fonctionnement, les deux décodeurs latéraux sont utilisés quand seul le canal qui leur est associé fonctionne. On suppose qu'il existe toujours au moins un canal en état de fonctionnement.

Le débit du canal i est noté R_i. En notant $\left\{\hat{X}_k^{(i)}\right\}_{k=1}^{N}$ la séquence reconstruite par le décodeur i, les performances du système pour un débit donné sont caractérisées par trois distorsions : la distorsion centrale D_0 qui est l'erreur de reconstruction au décodeur central, et les deux distorsions latérales, D_1 et D_2, aux décodeurs latéraux. Ces distorsions s'expriment par :

$$D_i = \frac{1}{N}\sum_{k=1}^{N} E\left[\delta_i\left(X_k, \hat{X}_k^{(i)}\right)\right], \quad i \in \{0,1,2\} \tag{3.1}$$

Où les δ_i sont des mesures de distorsion, à valeurs réelles, non négatives, et éventuellement distinctes.

La région débit-distorsion pour la source $\{X_k\}$ est définie comme étant l'ensemble des quintuplets $(R_1, R_2, D_0, D_1, D_2)$ atteignables au sens de Shannon. Comme le décodeur 1 recoit R_1 bits, sa distorsion ne peut être inferieure à $D(R_1)$ ou $D(.)$ est la fonction débit-distorsion de la source. En utilisant la même notation pour les deux autres décodeurs, on obtient les inégalités suivantes :

$$D_0 \geq D(R_1 + R_2) \tag{3.2}$$

$$D_1 \geq D(R_1) \tag{3.3}$$

$$D_2 \geq D(R_2) \tag{3.4}$$

Une égalité parfaite pour (3.2), (3.3) et (3.4) signifierait un débit optimal $R_1 + R_2$ pour la description centrale qui se diviserait en deux débits optimaux R_1 et R_2 pour les descriptions latérales. Malheureusement, il n'est généralement pas possible d'atteindre cet état idéal car des descriptions optimales individuellement sont souvent très similaires et donc redondantes lorsqu'elles sont combinées.

Le compromis fondamental de ce type de problème réside dans la construction de descriptions suffisamment bonnes individuellement tout en étant suffisamment différentes [136].

III.1.3. Codage à N descriptions :

Le problème de codage *MD* peut être généralisé à plus de deux canaux et plus de trois récepteurs. L'extension naturelle est N descriptions avec $2N - 1$ décodeurs. Cette généralisation a été étudiée par Witsenhausen [77] pour le cas où la source a un débit entropique fini et où une communication sans perte est requise quelque soit le nombre de descriptions reçues. En normalisant le débit de la source et en supposant une utilisation identique de chaque canal, ce dernier doit alors recevoir un débit de $1/(N - k)$, ou k est le nombre de descriptions reçues. Cette limite est atteinte en utilisant des codes de Reed-Solomon tronqués. Des résultats similaires sont présentés dans un cadre plus général dans [78] et une situation avec trois canaux et sept décodeurs a été étudiée par Zhang et Berger dans [79].

Plus récemment, Venkataramani, Kramer et Goyal [80] ont proposé une généralisation des résultats obtenus par El Gamal et Cover [71] pour $N = 2$ descriptions. Ils ont défini une borne externe de la région débit-distorsion pour une source gaussienne et une mesure de distorsion quadratique et ont montré que cette borne correspond à la région des débits atteignables seulement lorsque l'on s'intéresse aux performances du décodeur central et des N décodeurs à description simple (codage *SD*). Puri, Pradhan et Ramchandran [81][82] sont également parvenus à caractériser les performances du problème du codage à N descriptions. Pour ce faire, ils ont appliqué les résultats présentés dans [48] au problème de codage symétrique à N descriptions en utilisant la structure de codage proposée par Albanese et al [83].

Les techniques utilisées pour trouver cette région de débit sont inspirées de celles utilisées dans la démonstration du théorème de Wyner-Ziv [84] et sont basées sur une philosophie de codage avec information adjacente.

L'utilisation de ces principes est intéressante car elle contraste avec les approches basées sur une structure de raffinement progressif utilisées dans [71][79].

Leur approche sera décrite pour $N = 3$ descriptions mais est généralisable au cas $N > 3$. Soit X une source ayant pour densité de probabilité la fonction $q(x)$. X est segmentée en 3 couches $Y_1, Y_2\ et\ Y_3$. Soient : Y_3 l'alphabet de Y_3 et Y_{ij} l'alphabet de Y_{ij} où :

$$i \in I, I = \{1,2\} \text{ et } j \in J, J = \{1,2,3\}.$$

Le débit total est partitionné en trois débits : $R = R_1 + R_2 + R_3$. La couche de base est codée par un code correcteur d'erreurs $(3,1)$ en trois variables Y_{11}, Y_{12}, Y_{13} pour un débit R_1. La réception de deux descriptions doit être suffisante pour décoder la deuxième couche. On fait l'hypothèse que la transmission de Y_2 utilise le paradigme distribué. La réception de deux descriptions permet de recevoir une des paires $(Y_{11}, Y_{12}), (Y_{12}, Y_{13})\ ou\ (Y_{13}, Y_{11})$. Chacune de ces paires peut être traitée comme information adjacente pour estimer X. Grâce au codage de source avec information adjacente, une seule couche de raffinement est alors nécessaire [136].

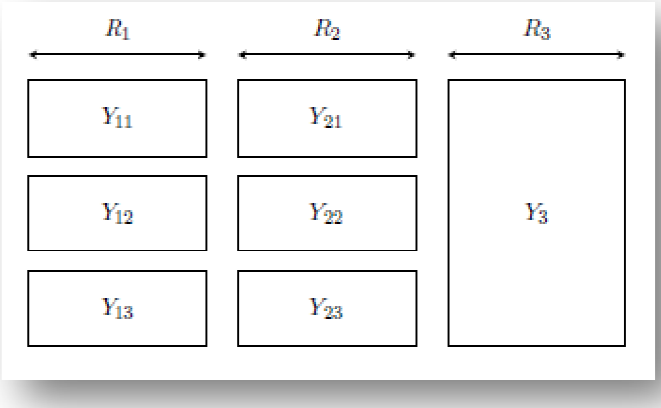


Figure 3.2. Codage de canal par descriptions multiples utilisant une concaténation de codes correcteur $(3,1), (3,2)\ et\ (1,3)$.

III.1.4. Codage à descriptions multiples basé sur la quantification :

III.1.4.1. Quantification scalaire à descriptions multiples :

Les premiers résultats en codage *MD* ont été obtenus par la technique de quantification scalaire à descriptions multiples introduite par Vaishampayan. La *MDSQ* est utilisée essentiellement pour le codage de deux descriptions et est constituée de deux parties :
Un quantificateur scalaire, et une assignation d'index qui sépare l'information de chaque échantillon quantifié en deux descriptions complémentaires du même échantillon. Un exemple est illustré sur la figure 3.3.

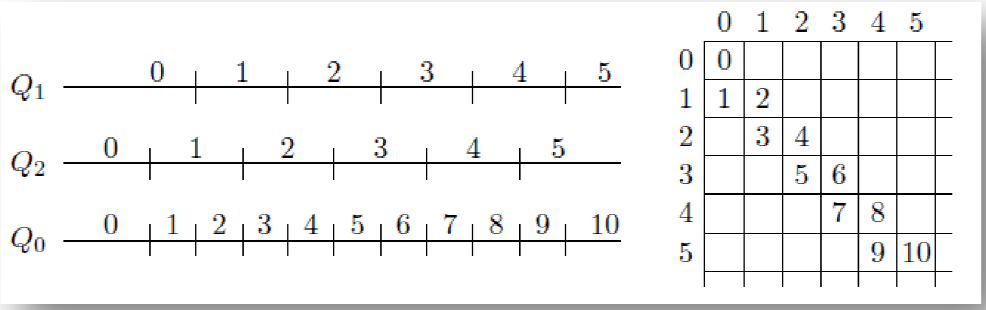


Figure 3.3. Exemple de *MDSQ* basée sur une indexation entrelacée modifiée décrite dans [85].

Le décodeur central reçoit quant à lui les deux informations, $Q_1(x) = k_1$ et $Q_2(x) = k_2$. Dans l'exemple, les cellules d'intersection sont deux fois plus petites que celles des quantificateurs individuels, la distorsion centrale est alors quatre fois plus petite que la distorsion latérale.

Asymptotiquement, si les débits latéraux sont égaux $R_1 = R_2 = R$, alors les trois distorsions D_0, D_1 et D_2 sont toutes de l'ordre de $O(2^{-2R})$. Si la distorsion centrale D_0 dépend essentiellement du nombre de cellules, la distorsion latérale dépend surtout de l'indexation de chaque quantificateur : selon cette indexation, l'étalement des cellules peut beaucoup varier.

Vaishampayan donne alors dans [85] un algorithme pour remplir les tables d'index de façon à obtenir des résultats optimaux pour les distorsions centrales et latérales. Par la suite, Berger et Reingold [86] ont généralisé la construction de matrices d'assignation d'index pour N descriptions.

Dans le but d'améliorer les performances du système, on peut malgré tout souhaiter utiliser un code de longueur variable, tel qu'un code de Huffman ou un code arithmétique. Un tel codeur est appelé quantificateur scalaire à descriptions multiples avec codage entropique et nous pouvons trouver un exemple d'étude sur ce sujet dans [87].

III.1.4.2. Quantification vectorielle à descriptions multiples :

Nous avons vu la construction de descriptions avec une quantification scalaire. Il est également possible de créer des descriptions avec une quantification vectorielle [88]. Ce problème, appelé quantification vectorielle sur réseau de points régulier à descriptions multiples (*MDLVQ)* a été particulièrement étudié pour deux descriptions équilibrées par Servetto, Vaishampayan et Sloan [89][90]. Il revient à assigner des paires d'indices aux différents points d'un dictionnaire de quantificateur vectoriel.

La source $\{X_n\}$ est divisée en vecteurs de longueur L. Les vecteurs sont ensuite quantifiés par un treillis fin Λ. Á chaque point sur Λ est assigné une unique paire de points $(\Lambda'_1, \Lambda'_2) \in \Lambda' \times \Lambda'$ par une fonction d'indexation vectorielle l.

Toute la difficulté de la *MDLVQ* consiste à trouver la fonction d'indexation qui donne des résultats optimaux pour un treillis donné. Dans [90], les auteurs proposent d'exploiter les symétries des réseaux de points pour construire des sous-treillis optimaux.

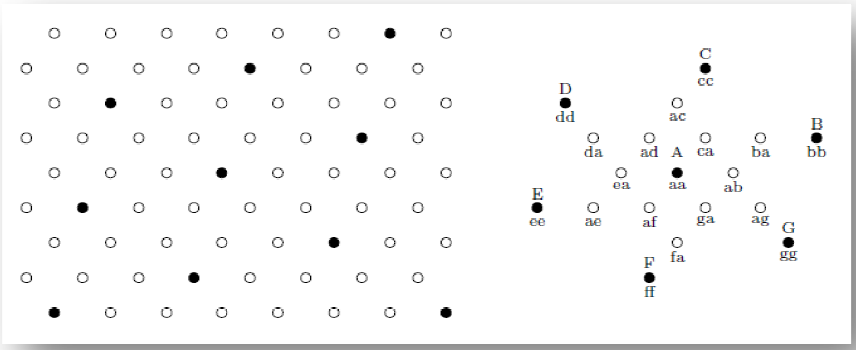


Figure 3.4. Exemple de treillis pour la *MDLVQ* : le treillis hexagonal (gauche), et l'indexation optimale associée (droite).

Considérons par exemple le treillis hexagonal représenté sur la figure 3.4. Tous les points appartiennent au treillis Λ mais seuls les points noirs appartiennent au treillis Λ'. Si par exemple la paire codée est le point ab, le symbole reconstruit au premier décodeur latéral sera le point A, et celui du deuxième décodeur sera le point B [136].

III.1.5. Codage à descriptions multiples basé sur la transformation :

Les techniques de codage de source se servent de transformations orthogonales tels que la transformée en cosinus discrète ou la transformée en ondelettes discrète pour décorréler le signal et en obtenir une représentation compacte qui supporte la suppression d'une partie de l'information au prix d'une perte de qualité acceptable lors de la reconstruction.

Cette opération contribue à diminuer la robustesse sur des canaux bruités, car en effet, si un symbole est perdu, les symboles voisins décorrélés ne permettront pas d'en faire une estimation correcte. Deux approches sont possibles pour introduire de la corrélation dans le domaine transformée. La première consiste à remplacer la transformation classique par une transformation qui permet à la fois de générer des descriptions et de contrôler leur niveau de redondance.

Par exemple, la *DCT* peut être remplacée par une transformée orthogonale recouvrante *LOT* (Lapped Orthogonal Transform) [91] suivie d'une répartition des blocs dans des descriptions [92]. De son coté, la *DWT* peut être remplacée par des bancs de filtres multi-ondelettes [93][94] dont les sorties sont les descriptions [95]. La deuxième approche consiste à ajouter une transformation linéaire entre la transformation classique et la quantification de façon à introduire de la corrélation entre les symboles indépendants.

Le *MDTC*, a été introduit par Wang, Orchard, Reibman et Vaishampayan [96][97] pour le cas de deux variables et a été généralisé par Goyal et Kovacevic pour n variables et une plus grande classe de transformées [98], et en introduisant l'utilisation de la théorie des expansions sur bases de fonctions redondantes [99]. Ils considèrent le cas de deux variables : $x = (x_1, x_2)$ indépendantes ayant comme variance σ_1^2 et σ_2^2, respectivement. Le principe du *MDTC* consiste à transformer linéairement ces deux variables non corrélées en deux variables $y_1$et $y_2$corrélées. Les paramètres de la transformée contrôlent la corrélation entre y_1 et y_2 qui, à son tour, contrôle la redondance du codeur *MDTC*. Le schéma de codage de base est illustré sur la figure 3.5.

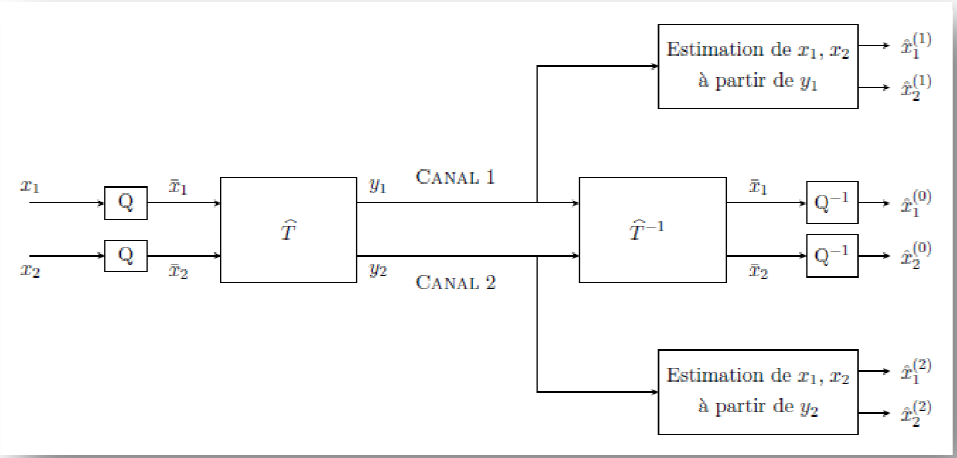


Figure 3.5. Schéma *MDTC* pour une paire de variables gaussiennes indépendantes.

Comme la transformée est, en général, non orthogonale, la quantification de y_1 et y_2 donnerait de très mauvaises performances. C'est pourquoi, x_1 et x_2 sont d'abord quantifiées en $\bar{x}_1$ et $\bar{x}_2$ avant d'être transformées par $\hat{T}$, une version discrète d'une transformée linéaire continue et inversible T.

On peut montrer que si T est une transformée linéaire continue de déterminant 1, elle est inversible. La première étape pour discrétiser une transformée linéaire consiste à la factoriser en un produit de matrices triangulaires inferieures et supérieures à diagonale unité $T = T_1 T_2 \dots . T_k$. La version discrète est alors donnée par :

$$\hat{T}(x_q) = \left[T_1 \left[T_2 \dots \left[T_k x_q\right]_\Delta\right]_\Delta\right]_\Delta \tag{3.5}$$

où Δ est le pas du quantificateur utilisé par le codeur.

Au récepteur, si les deux séquences y_1 et y_2 sont disponibles, la transformée discrète est inversée pour obtenir $\bar{x}_1$ et $\bar{x}_2$ et une quantification inverse est appliquée pour obtenir $\hat{x}_1^{(0)}$et $\hat{x}_2^{(0)}$. La qualité de la reconstruction ne dépend alors que de la quantification initiale. Dans le cas où une seule description est reçue, par exemple y_1, y_2 sont estimées à partir des composantes reçues en utilisant la corrélation statistique introduite par la transformée $\hat{T}$ [136].

III.1.6. Codage à descriptions multiples basé sur une décomposition en sous-bandes :

Cette technique a été proposée simultanément par Dragotti, Servetto et Vetterli [100] et Yang et Ramchandran [101] dans le cas de bancs de filtres à deux canaux.

Elle repose sur l'utilisation d'une décomposition en sous-bandes, ou d'une *DWT* souvent utilisée dans les systèmes de codage d'image ou de vidéo. L'utilisation de bancs de filtres permet d'introduire de la corrélation entre deux descriptions $y_1[n]$et $y_2[n]$. L'approche proposée dans [101] pour synthétiser les filtres consiste à minimiser la distorsion obtenue avec une seule description pour un niveau de redondance donnée.

L'optimisation repose sur une formulation Lagrangienne qui est résolue directement dans le domaine fréquentiel pour produire les réponses des filtres. Si un canal est défaillant, par exemple le canal 1, les symboles manquants $y_1[n]$ sont estimés linéairement à partir des symboles présents au décodeur $y_2[n]$.

Le codage par sous-bandes est peu utilisé en pratique pour produire des descriptions multiples car il nécessite de synthétiser un banc de filtres différent pour chaque niveau de redondance, ce qui limite sa capacité d'adaptation à des conditions de transmission variables.

III.1.7. Codage à descriptions multiples basé sur les codes correcteurs d'erreurs :

Les méthodes vues jusqu'à présent introduisaient de la redondance au niveau du codage de source. La méthode présentée dans [83] et [102] propose d'utiliser des techniques de codage de canal pour ajouter de la redondance dans le signal à transmettre.

On suppose pour cela que le flux d'information à coder est organisé de façon hiérarchique et segmenté en couches d'importances décroissantes. Nous pouvons alors protéger ces couches par des codes correcteurs d'erreurs de redondance également décroissante. Cette idée est connue sous le nom de protection inégale contre les erreurs (Unequal Error Protection). Cette technique peut être directement décrite dans le cas général de N descriptions. L'application visée est la transmission par paquets sur un réseau ou les données à transmettre sont représentées par un flux binaire hiérarchique progressif. Le mécanisme de paquetisation du flux est illustré sur la figure 3.6. La première étape consiste à segmenter le flux hiérarchique en N couches de résolution croissante. La $i-éme$ couche doit être décodable lorsque i descriptions parviennent au décodeur, et ce quelles que soient les descriptions reçues. Pour parvenir à cela, on divise la $i-éme$ couche en i parties égales, que l'on place dans les i premiers paquets. Pour compléter les autres paquets, on utilise des codes correcteurs d'erreurs en bloc (FEC) de paramètres $(N, i, N-i+1)$. Ces codes transforment i bits source en N bits codes par l'ajout de $N - i$ bits de redondance.

Cette redondance permet de corriger $N - i$ bits erronés. Ces blocs sont appliqués verticalement aux i tronçons de la $i - eme$ couche.
Grâce à ce mécanisme, chaque description contient une partie de chacune des N couches du flux initial, avec la propriété que la $i - éme$ couche peut être récupérée à partir de i paquets quelconques.

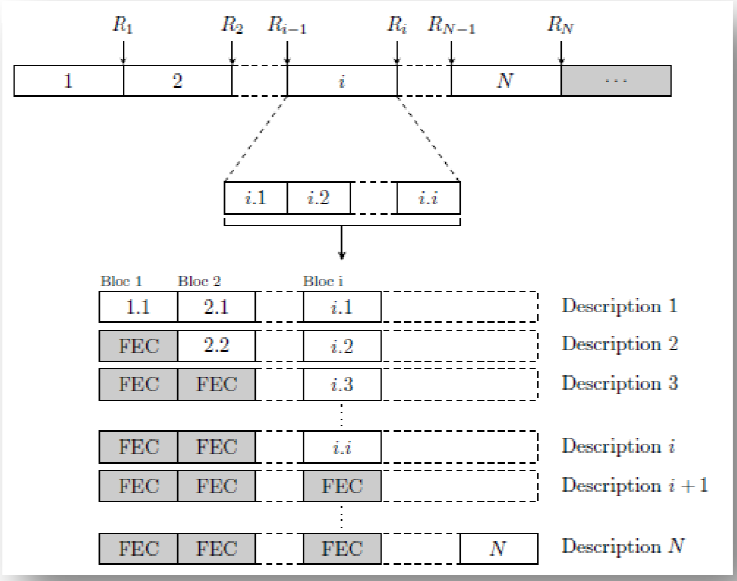


Figure 3.6. Schéma de paquetisation d'un flux hiérarchique binaire [136].

III.1.8. Codage vidéo à descriptions multiples :

Une des applications du codage *MD* qui a reçu le plus d'attention récemment est la transmission vidéo. La motivation principale pour l'utilisation du codage *MD* dans ce domaine est sa capacité d'assurer une qualité minimum sans nécessiter la retransmission de paquets perdus.

Cette particularité rend le codage *MD* intéressant pour des applications interactives, en temps réel, pour lesquelles une retransmission de paquets serait jugée inacceptable. Le problème posé par le codage vidéo par des descriptions multiples est lié à la prédiction de mouvement.

La prédiction de mouvement est un composant fondamental dans tous les codeurs vidéo standard. Cependant, du fait de la prédiction, une erreur de transmission sur une description peut se répercuter sur un nombre important d'images, on parle alors de drift. Il est donc impératif que les codeurs vidéo par descriptions multiples soient construits en tenant compte de l'erreur de propagation que des erreurs de transmission peuvent engendrer.

Wang et al [103] ont proposé une classification des codeurs vidéo par descriptions multiples suivant le fonctionnement de leur prédicteur et suivant le type de redondance utilisé. Les prédicteurs sont rangés en trois classes :

- ***Classe A*** ne souffre pas de problème de prédiction. Ainsi, l'encodeur et le décodeur utilisent les mêmes signaux de prédiction.

- ***Classe B*** est identique à celui qui serait utilisé pour le codage *SD*. Ce prédicteur minimise l'erreur de prédiction et donc n'introduit aucune redondance.

- ***Classe C*** peut contrôler le compromis entre l'efficacité de la prédiction et l'introduction de drift.

Le codage *MD* a également été utilisé en codage vidéo $t+2D$ avec filtrage temporel compensé en mouvement. McCanne et al [104][105] ont proposé l'utilisation du codage conjoint source-canal pour la transmission multipoints de vidéo sur un réseau hétérogène. Dans leur approche, chaque récepteur peut choisir dynamiquement la capacité du réseau local en ajustant la qualité de la vidéo reçue.

Sur un réseau tel qu'Internet, plusieurs descriptions peuvent être envoyées à un récepteur suivant plusieurs chemins [106]. Dans [107], Pereira et al ont proposé un schéma de codage *MD* basé sur une *DWT* et une technique d'allocation de débit dérivée de [108] et adaptée au codeur vidéo *3D* présenté dans [109] [136].

III.1.9. Codage à descriptions multiples utilisant un pré et post traitement:

III.1.9.1. Codage d'images par descriptions multiples utilisant un pré et post traitement :

Dans cette approche *MDC*, l'étape la plus importante est de séparer des lignes et colonnes des images (sur échantillonnée) en sous bandes paires et impaires pour constituer des sous images (descriptions). Les pixels des lignes paires et colonnes impaires sont attribués à une description. Tandis que et les lignes impaires et colonnes paires sont attribuées à une deuxième description. Ces images sont alors codées en deux descriptions utilisant un standard de compression tel que *JPEG* (*JPEG*2000). En sur-échantillonnant l'image originale, des échantillons additionnels sont introduits entre les échantillons originaux.

Cette méthode de sur-échantillonnage consiste à introduire des coefficients nuls dans le domaine fréquentiel (*DCT*). L'image sur-échantillonnée résultera de la *DCT* inverse de l'opération précédente. L'image d'entrée est transformée avec *DCT-2D*, et est ensuite emplit dans les deux dimensions. L'image résultante est inversée (*IDCT*) produisant une image sur-échantillonnée. Par conséquent les valeurs de pixels de la nouvelle image obtenue sont plus corrélées que les pixels de l'image originale. Les deux descriptions obtenues après le sous échantillonnage de l'image corrélée sont codées par un codeur comme *JPEG*. Á la réception, le décodeur décode simplement les paquets qu'il reçoit. C'est l'étape de post traitement qui fait l'estimation des pixels manquant et recombine les sous images. Le schéma de la méthode proposée est illustré sur la figure.3.7.

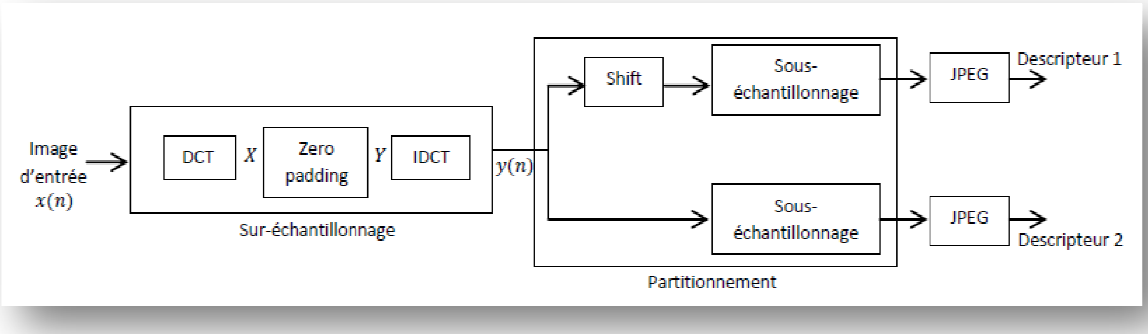


Figure 3.7. Processus de pré-traitement proposé pour générer deux descriptions.

Y est le signal obtenu par l'ajout de $M - N$ zéros au signal X. Une transformée *DCT* inverse est appliquée à Y pour obtenir l'image sur-échantillonnée y. Le signal d'entrée $x(n)$ peut être reconstruit à partir du signal $\tilde{y}(n)$ qui correspond à N points du signal $y(n)$. Le signal $\tilde{y}(n)$ est réparti en deux descriptions, l'une portant les échantillons pairs du signal et l'autre contenant les échantillons impairs. Si une description est perdue, $x(n)$ sera reconstruite à partir de $M/2$ échantillons de $y(n)$ reçus.

Pour estimer une description perdue, ses pixels seront remplacés par les plus proches de la même colonne de la description reçue. Les résultats de l'application de ce schéma *MDC* avec un codec *JPEG*, considérant le cas de deux descriptions, ont démontré sa performance par rapport à l'approche décrite dans [73] et par rapport à un codage à descriptions multiples scalaire. Par comparaison de la qualité subjective de l'image reconstruite à partir d'une seule description, sans sur-échantillonnage, présente une certaine distorsion sur les bords de l'image.

III.1.9.2. Codage vidéo par descriptions multiples utilisant un pré et post traitement :

Dans [74] une approche de codage vidéo par descriptions multiples consiste à sur-échantillonner la vidéo et la décomposer en k sous-images par un système de prétraitement. Ces sous images (descriptions) sont codées et transmises indépendamment. La vidéo reconstruite au niveau du récepteur est générée par une étape de post-traitement. Les deux descriptions ainsi obtenues sont corrélées en appliquant un sur-échantillonnage des images d'entrée qui sont séparées ensuite pour former les deux séquences vidéo à coder et transmettre indépendamment sur différents canaux. Le même principe de séparation des pixels en lignes et colonnes paires et impaires a été adopté dans le domaine spatiale pour former les séquences vidéos à transmettre [111]. Selon les résultats obtenus en terme de *PSNR* de la vidéo reconstruite en fonction du débit ont montré une amélioration lorsque le sur-échantillonnage a augmenté. Plusieurs schémas de codage vidéo à descriptions multiples sont présentés avec un pré et post traitement [72], la redondance est introduite en ajoutant des zéros et les descriptions sont générées en sur-échantillonnant les images obtenues. Bai et al [111] ont proposé un schéma *MD* pour le codage vidéo illustré sur la figure 3.8.

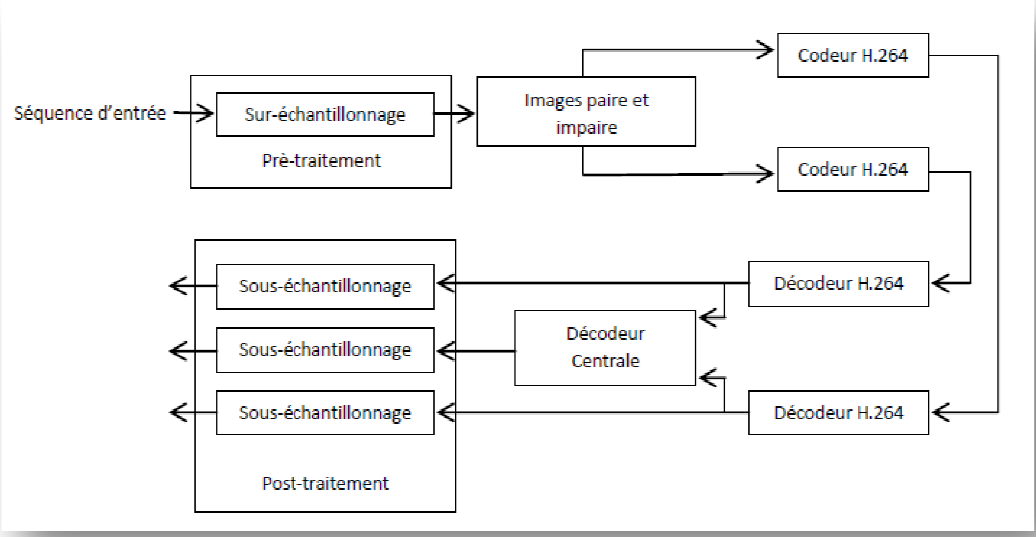


Figure 3.8. Schéma de codage vidéo *MD* proposé dans [111].

Dans ce schéma le taux de sur-échantillonnage diffère selon le mouvement de l'information entre les images adjacentes, d'autre part, si le mouvement de l'information entre les images est lisse moins d'images de redondance sont nécessaires. L'objectif de telle méthode de sur-échantillonnage est de générer des descriptions où le mouvement est lisse afin de mieux estimer les images perdues à la reconstruction.

Les vecteurs de mouvement sont calculés et le maximum est donné par $\|MV\| = \sqrt{x^2 + y^2}$, x et y sont les coordonnés du vecteur de mouvement maximal qui sera comparé à un seuil T. Les images paires sont insérées pour que le nombre d'images soit égal dans les deux descriptions, à la fin du prétraitement une étiquette avec un bit '1' ou '0' est attribuée à chaque image pour distinguer la trame originale de la trame extra. À l'étape de post traitement, lorsque les deux descriptions sont reçues, les trames sont réarrangées d'après les étiquettes '1' et '0' et la vidéo reconstruite pourra être sous-échantillonnée afin d'obtenir la vidéo reconstruite finale. Les résultats présentés dans ce travail ont été comparés à un schéma similaire sans sur-échantillonnage des séquences d'entrée.

III.2. Représentation parcimonieuse :

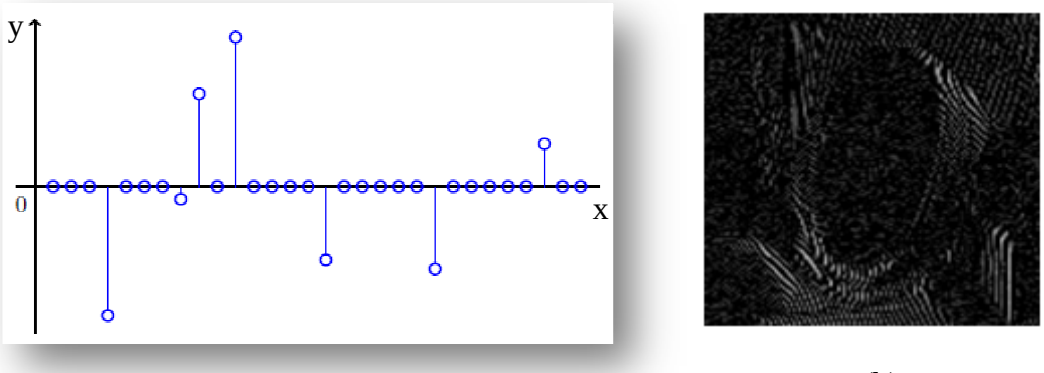
III.2.1 Introduction :

La représentation parcimonieuse est devenue un sujet de recherche très attractif au cours des dernières années. Beaucoup de nouveaux algorithmes ont été développés qui tiennent d'avantage la représentation parcimonieuse pour une large gamme d'applications de traitement des images tels que : L'inpainting, le débruitage, et la compression en raison de la nature compressible de cette représentation. Un schéma de représentation sparse se compose généralement d'un dictionnaire et un algorithme qui sélectionne l'atome et les coefficients de pondération afin de produire une approximation linéaire du signal d'entrée. Dans cette section, on commence par présenter le problème de parcimonie d'une façon générale, puis donner une formulation générale sur quelques algorithmes de représentation parcimonieuse.

III.2.2. Représentation parcimonieuse d'un signal :

Les signaux peuvent être modélisés de façon continue ou de façon discrète. On s'intéresse ici au cas de signaux discrets et de longueur finie, qui ont en pratique *n* échantillons. Il est courant de vouloir approximer de tels signaux via une classe de signaux plus facilement manipulables. C'est le cas des signaux dits parcimonieux. les représentations parcimonieuses consistent en la décomposition du signal sur un dictionnaire comprenant un nombre d'éléments (atomes) supérieur à la dimension du signal. Cette décomposition introduit dans le nouveau signal un grand nombre de valeurs nulles, d'oú vient la représentation parcimonieuse.

On trouve aussi les termes représentation "éparse" et "creuse". Au sens strict, un signal est dit parcimonieux lorsqu'il ne comporte qu'un faible nombre de coefficients non nuls face à la longueur de sa représentation, comme le signal de la figure 3.9. La connaissance des valeurs et positions de ces quelques coefficients suffisent pour définir un tel signal. La parcimonie a des applications directes en compression de signaux et des images.



(a) (b)

Figure 3.9. Exemple d'un signal parcimonieux (a) Signal 1D, (b) Signal 2D.

Depuis quelques années, les représentations parcimonieuses de signaux et d'images font l'objet de développements importants [112][113][114]. Le développement de ces techniques est issu des travaux de Stéphane Mallat et al [115] et F. Bergeaud et al [116]. Il existe trois problématiques que posent ces représentations :

La définition du critère de parcimonie, méthodes de calcul de x, connaissant le signal y réelle et le dictionnaire D.

Notant :

$$y = Dx \tag{3.6}$$

On peut cependant se demander si, le signal donné y, on peut trouver un signal parcimonieux x qui soit proche de y. Pour cela, on se munit d'une mesure de distorsion $(x, y) \rightarrow d(x\ y) \in \mathbb{R}^+$. Elle sert à mesurer l'erreur que l'on engendre en approximant y par x. Elle augmente lorsqu'on s'éloigne du signal d'origine et vaut 0 uniquement dans le cas où $x = y$. Une mesure couramment utilisée est la mesure de distorsion ℓ_2 (erreur quadratique) :

$$d(x\ y) = \|y - x\|_2^2 \tag{3.7}$$

III.2.3. Dictionnaire :

Dans l'analyse des composantes parcimonieuses on appelle dictionnaire $D \in \mathbb{R}^{N \times M}$ la matrice contenant l'ensemble des fonctions prototypes avec lesquelles on modélise le signal, et chacune de ces atomes est une colonne du dictionnaire. Si on veut que le signal que l'on observe soit bien représenté, il faut un dictionnaire qui s'en soit adapté. Cependant, les signaux réels étant très variés il faudrait construire un dictionnaire long et le calcul serait difficile en un temps raisonnable.

De nombreuses transformées ont été développées qui permettent d'approximer les images, comme les ondelettes ou les curvelets, et forment des dictionnaires à transformée et reconstruction rapide. La génération d'un dictionnaire pour les applications de traitement d'images peut être réalisée soit à partir d'un dictionnaire pré-calculé correspondant à une famille de fonctions connues : ondelettes, cosinus discret, soit à partir des images elles mêmes, ce qui permet une adaptation très forte de la représentation parcimonieuse aux images considérées.

III.2.4. Les algorithmes d'approximation parcimonieuse :

Un signal $y \in R^N$ composé de N échantillons et une matrice $D \in \mathbb{R}^{N \times M}$ composée de M colonne appelées "atomes" $\{D_m\}_{m=1}^{M}$, la décomposition du signal y est donnée par :

$$y = Dx + r \quad (3,8)$$

$$= \sum_{m=1}^{M} D_m x_m + r \quad (3,9)$$

Où $x \in R^M$ est le vecteur de coefficients et r l'erreur résiduelle.
Le vecteur signal qui peut être représenté d'une manière parcimonieuse avec une erreur acceptable est appelé compressible. Comme la compressibilité du signal dépend du dictionnaire utilisé pour obtenir la représentation, divers auteurs ont mis au point de nouveaux algorithmes qui produisent des dictionnaires adaptés à une classe de signal particulier.
Les dictionnaires formés sont devenus un élément important des systèmes de représentation parcimonieuse, c'est ce qui a aidé plusieurs auteurs à avoir de bons résultats.

Il existe de nombreuses méthodes utilisées pour résoudre le problème d'approximation parcimonieuse. Les deux méthodes les plus courantes utilisées sont : "Poursuite de base *BP* (Basis Pursuit)" et le "Poursuite adaptative *MP* (Matching Pursuit)".

III.2.4.1. :L'approche Poursuite de base BP:

Chen, Donoho et Saunders [117][118] ont développé une méthode appelée Basis Poursuit. L'idée du Poursuite de base est de remplacer le problème de parcimonie difficile avec un problème d'optimisation plus facile. Voici la définition formelle du problème :

$$\min \|x\|_0 \ tel\ que\ y = Dx \quad (3.10)$$

La difficulté avec le problème ci-dessus est la norme ℓ_0. L'algorithme poursuite de base consiste à remplacer la norme ℓ_0 avec ℓ_1 pour rendre le problème plus facile à travailler. L'algorithme résultant, connu comme poursuite de base est généralement exprimée comme la solution du problème suivant :

$$\min\|x\|_1 \ tel\ que\ y = Dx \qquad (3.11)$$

La solution consiste généralement en d vecteurs linéairement indépendants spécifiés par y.

III.2.4.2. L'approche Poursuite adaptative MP:

Le Matching Pursuit introduit par Mallat et Zheng [115] est un algorithme itératif, de conception assez simple, qui recherche des approximations de plus en plus précises de la solution par optimisations unidimensionnelles. Il s'agit d'un algorithme glouton qui, au lieu de chercher la solution globale recherche successivement la solution optimale d'un sous-problème, à savoir la meilleure approximation du signal.

Rappelons ici que nous considérons le problème général de la décomposition d'un signal y sur un dictionnaire D constitué d'un ensemble de vecteurs unitaires $\{a_\gamma\}_{\gamma\in\Gamma}$.
Le nombre d'itérations n'est donc pas nécessairement égal au nombre total d'atomes a_γ sélectionnés pour la représentation x. L'intérêt de cet algorithme est sa simplicité d'implémentation.

On recherche par itération la solution optimale, la décomposition du signal x sur le dictionnaire aprés *MP* est donnée par :

$$y = \sum_{m=1}^{M} a_{km}x_m + r_m \qquad (3,12)$$

Oú r^0 le résidu initial est le signal x et à chaque itération l'atome a_{ki} sélectionné est celui dont le produit scalaire avec le résidu r_m le plus grand.

Voici décrit ci-après, les étapes du Matching Pursuit :

- La première approximation s'obtient en calculant le projeté orthogonal du signal observé y sur l'atome qui lui est le plus corrélé.

- Lors des itérations suivantes, on recherche l'atome le plus corrélé au résidu courant et on projette ce résidu sur l'atome sélectionné et

l'estimation obtenue par l'ajout de ce nouvel atome est ajoutée à l'estimation courante du résidu, pour former l'estimation courante du signal y.

III.2.4.3. L'approche Poursuite adaptative orthogonale OMP:

L'Orthogonal Matching Pursuit est une extension de l'algorithme *MP* et consiste à trouver dans le dictionnaire le vecteur le plus corrélé avec le signal. L'*OMP* se base sur le même principe que le *MP* : sélectionner itérativement les atomes les plus corrélés au signal pour tendre vers une approximation de la solution, la différence réside dans la mise à jour de l'approximation, et donc du résidu r.

$$x_i = arg\ min_x \|y - D_i x\|_2^2 \qquad (3,13)$$

Soit $D_i = [a_1 \dots a_i]$ le sous dictionnaire de D regroupant les i premiers atomes sélectionnés et $x_i = [x_1 \dots x_i]$ le vecteur de coefficients correspondant.

L'algorithme matching pursuit orthogonal suit ces étapes :

- ***Initialisation :*** $r^{(0)} = x, x^{(0)} = 0$.
- ***Itération :*** Le critère d'arrêt est atteint en cherchant l'élément optimal et la mise à jour de l'approximation au rang k.

OMP permet de pallier au défaut du *MP* qui réside dans la possibilité de sélectionner plusieurs fois le même atome. Ces algorithmes de type glouton (*MP, OMP*) offrent de bonnes propriétés de décroissance de l'erreur entre le signal original et le signal parcimonieux. La famille des algorithmes gloutons est large et il en existe d'autres permettant notamment la sélection de plusieurs atomes par itération (*StOMP, ROMP, CoSaMP*). Les travaux de D. Needell [119] en présentent une étude approfondie.

III.2.5. Conclusion :

Dans cette partie, nous avons donné un bref aperçu des travaux existants dans le domaine de codage par descriptions multiples. Après avoir discuté brièvement le contexte historique et les notions générales du MDC, nous avons présenté les directions principales dans ce domaine: les méthodes MDC par quantification et les méthodes MDC par corrélation, MDC basé sur les codes correcteurs d'erreurs et MDC basé sur la transformation. Nous avons poursuivi par un rappel sur le codage vidéo à descriptions multiples dans la littérature. Finalement, nous avons présenté un bref état de l'art sur les schémas MDC basé sur le pré et post traitement pour l'image fixe et la vidéo.Nous avons vu aussi la notion de parcimonie. La parcimonie est présente partout pour différents types de signaux, comme les signaux audio ont une décomposition parcimonieuse dans le domaine temps-fréquence. Une représentation parcimonieuse est ainsi obtenue en décomposant une image courante dans une base d'ondelettes. L'intégration de la parcimonie peut se faire de différentes façons, par une approche sous optimale (MP, OMP) ou une approche globale (BP, BPD). De nos jours, le volume des données croît plus rapidement que les capacités de stockage, et beaucoup plus vite que la bande passante nécessaire à leur transmission. Le but de la compression d'un signal est de réduire sa dimension. La qualité de compression dépend du choix de la base orthogonale. Une équivalence entre la parcimonie du signal et l'efficacité de la compression de celui-ci est montrée dans le théorème de Donoho [120]. Nous nous intéressons dans la suite aux contributions de la parcimonie à la compression de l'image fixe puisqu'elle joue un rôle important dans différents domaines et conduit à des meilleures performances.

IV.1. Introduction :

L'objectif du codage à descriptions multiples est de reconstruire un signal source en plusieurs qualités à partir de différentes descriptions reçues. Dans les dernières années le codage *MD* a été largement appliqué pour la transmission des images fixes et la vidéo. Comme on a décrit précédemment, deux approche *MDC* ont été pratiquement appliquées le *MDSQ* [121] et le *MDTC*, cette dernière qui a été introduite par Wang et al [122][123] utilisant une transformée de corrélation *PCT* qui a pour rôle d'introduire une quantité de corrélation entre les descriptions. Ce schéma a été généralisé par Goyal et al [124] pour coder plus de deux descriptions.

Des techniques *MDC* basées sur la décomposition en ondelettes ont été pratiquement étudié [108][125] . En 2007, Khelil et al [126] ont proposé un schéma *MDC* basé sur une décomposition *DWT* utilisant la transformée de corrélation *PCT* entre les descriptions à transmettre, ce schéma constituait une amélioration du codeur *MDTC* proposé dans [124] [127].

Nous nous sommes intéressé dans ce chapitre à présenter nos résultats expérimentaux de l'implémentation d'un codeur *MDTC* basé sur les transformées *CT* et *NSCT*, puis voir les performances d'évaluation de qualité en codage vidéo *MD* par pré et post traitement des images de la séquence utilisant le banc de filtre NSP de la *NSCT*.

VI.2. Première Partie

Codage des images fixes par descriptions multiples

IV.2.1. Transformée de corrélation :

la transformée de corrélation est une méthode pour introduire la redondance entre des vecteurs de données décorrélés. Wang et al [122] ont démontré que pour coder deux descriptions, où chacune est susceptible d'être perdue, il est suffisant d'appliquer la transformée de corrélation de la forme suivante:

$$T_a = \begin{bmatrix} a & 1/2a \\ -a & 1/2a \end{bmatrix} \qquad (4.1)$$

Dans le cas de transmission de quatre descriptions [124] la transformée *PCT* en cascade illustrée sur la figure 4.1 est donnée par :

$$T = \begin{bmatrix} T_c & 0 \\ 0 & T_c \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} T_a & 0 \\ 0 & T_b \end{bmatrix} \tag{4.2}$$

Les paramètres a, b et c sont déterminés à partir de la redondance introduite par la transformée *PCT* et les variances des composantes à transmettre.

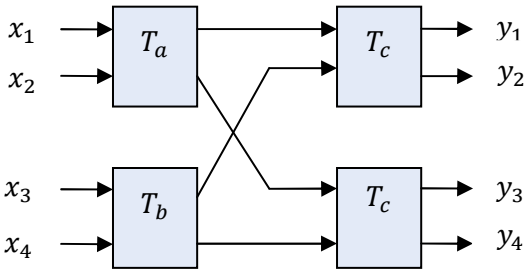


Figure 4.1. Structure de transformée en cascade pour un codeur *MDTC* à quatre descriptions.

IV.2.2. Paramètre d'évaluation :

Le critère d'évaluation le plus utilisé dans des schémas *MDC* est le rapport signal / bruit de crête $PSNR$ (Peak Signal to Noise Ratio) défini par :

$$PSNR(dB) = 10log_{10} \frac{max(X(i,j))}{EQM} \tag{4.3}$$

Où : $max(X(i,j))$ représente l'amplitude maximale de l'image originale. EQM représente la distorsion créée par la perte de l'information entre l'image originale et l'image décodée à la fin de la chaine de transmission.

Dans un schéma *MDTC*, la distorsion par description à un débit R est donnée par :

$$D = \frac{\pi e}{6} \sigma_1 \sigma_2 \, 2^{-2R} \tag{4.4}$$

σ_1, σ_2 représentent les variances des deux composantes.

IV.2.3. Comparaison Multirésolution pour le codage MDTC :

dans cette section on propose un schéma de comparaison *MDTC* [128] utilisant les transformées *DWT*, *CT* et *NSCT* pour décorréler l'image originale. Un schéma *MDTC* conventionnel suit ces étapes :

- ✓ Transformer l'image originale au domaine fréquentiel (*DWTI*, *CT* ou *NSCT*).
- ✓ Quantifier uniformément les coefficients résultants.
- ✓ Une transformée *PCT* est appliquée, les vecteurs obtenus sont alors corrélés.
- ✓ Les descriptions formées sont ensuite codées avec un codage entropique.

sachant que pour *DWT* les descriptions 1,2 et 3,4 forment respectivement deux paires de descriptions à transmettre à travers deux canaux différents. Ainsi pour la *CT* et *NSCT* les deux descriptions sont formées à partir des sous bandes issues du banc de filtre *LP* et *NSP* respectivement. La figure suivante représente un schéma *MDTC* à deux descriptions.

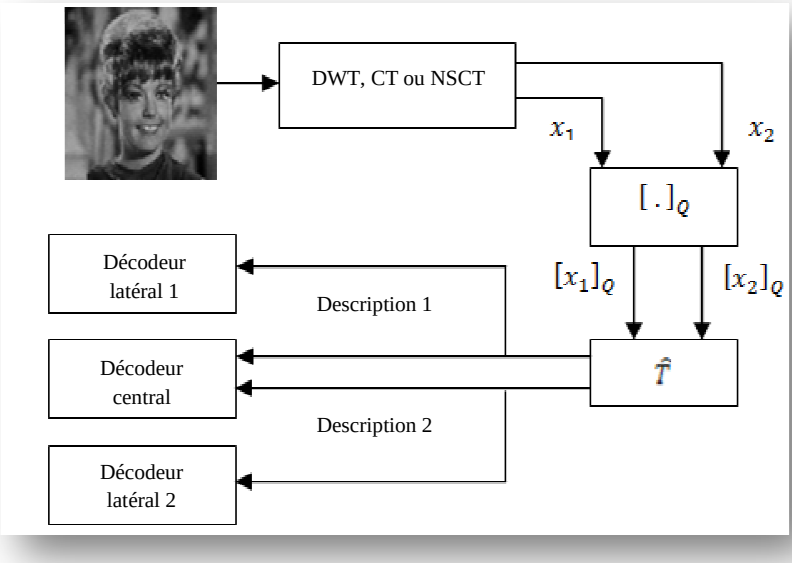


Figure 4.2. Schéma *MDTC* à deux descriptions.

Pour tester les performances de notre codeur *MDTC* basé sur la *CT* et *NSCT*, on a utilisé des images de taille 512×512 à différentes caractéristiques spatiales : $Lena, Barbara, Boat, Mandril, Zelda$ et $Livingroom$. Les résultats obtenus sont présentés en terme de $PSNR$ à un débit binaire $R = 2bps$.

	PSNR (dB) Central		
Images	DWT	CT	NSCT
Lena	32.05	34. 71	40.02
Barbara	30.29	33.04	35.53
Boat	32.69	34.91	39.41
Mandril	35.40	37.27	41.99
Livingroom	33.60	35.67	40.56
Zelda	35.02	37.08	43.29

Tableau 4.1. *PSNR*(dB) central pour *DWT*, *CT* et *NSCT*.

Les résultats en terme de qualité visuelle de l'image reconstruite présentés sur les figure 4.3-4.4 confirme l'efficacité des codeur *MDTC-CT* et *MDTC-NSCT* par rapport au schéma *MDTC-DWT* classique. Les résultats de simulation de la figure 4.5 indiquent que le *PSNR* central pour la *NSCT* et *CT* est meilleure que pour *DWT* dans un schéma *MDC*.

Figure 4.3. Image Lena (a) originale, Image reconstruite par (b) *DWT* (c) *NSCT* (d) *CT*.

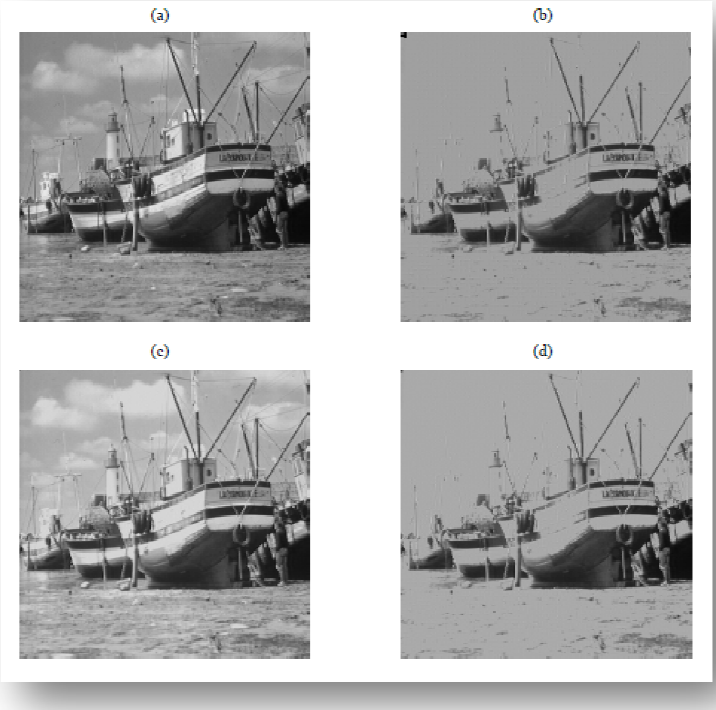


Figure 4.4. Image Boat **(**a) originale, Image reconstruite par (b) *DWT* (c) *NSCT* (d) *CT*.

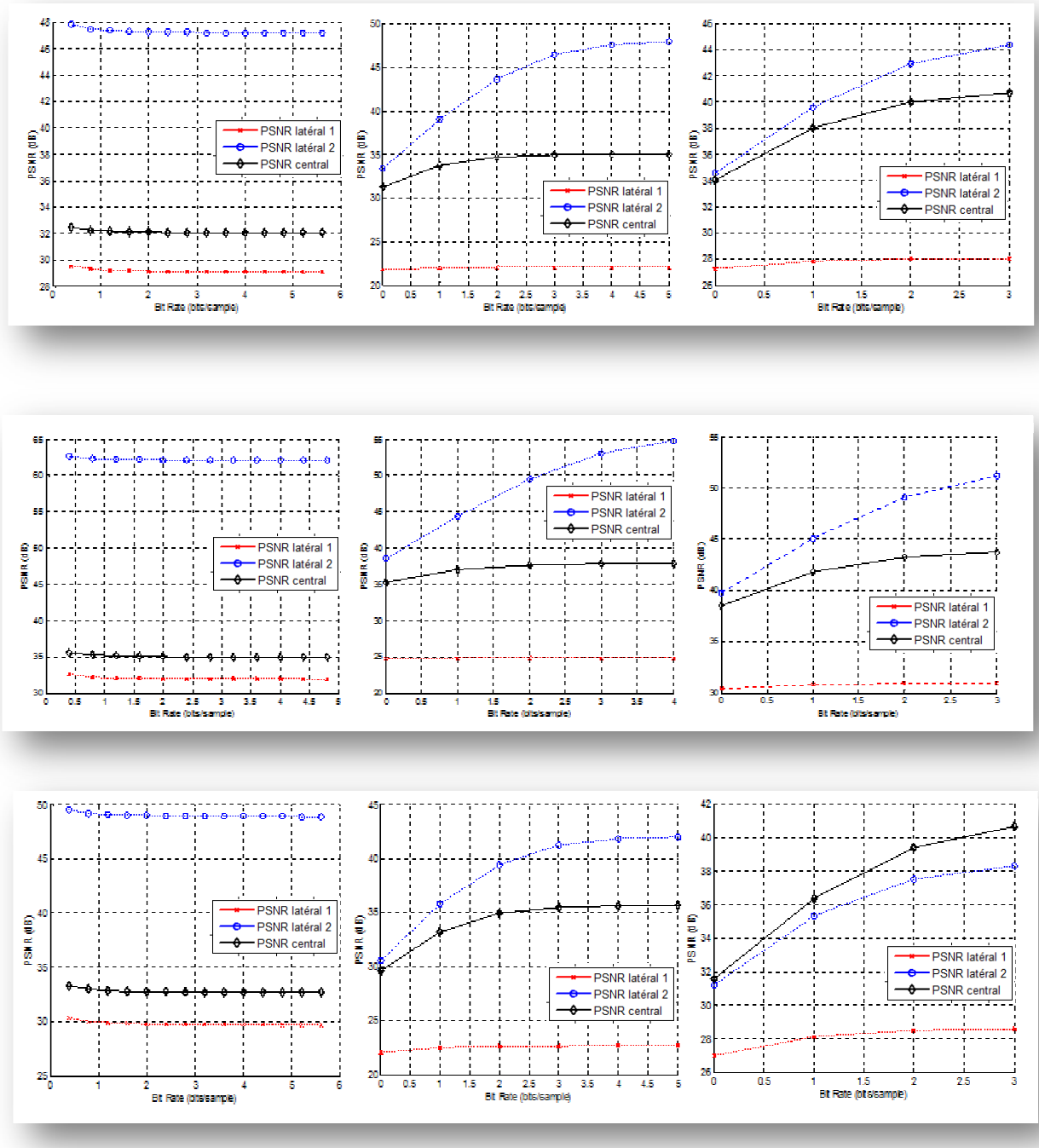


Figure 4.5. Valeurs de $PSNR$(dB) Central et latéraux en fonction du débit binaire utilisant les transformées DWT, CT et NSCT pour les images (a)Lena (b) Zelda (c) Boat.

IV.2.4. Codage MDTC basé sur la transformée en Contourlets :

Le schéma *MDTC-CT* présenté dans la section précédente est généralisé pour transmettre quatre descriptions sur différents canaux [129]. la figure 4.6 représente les blocs du codeur proposé.

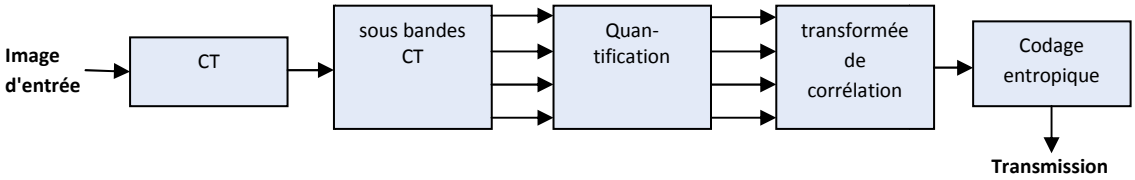


Figure 4.6. Schéma bloc d'un codeur *MDTC* basé sur la *CT* et *PCT*.

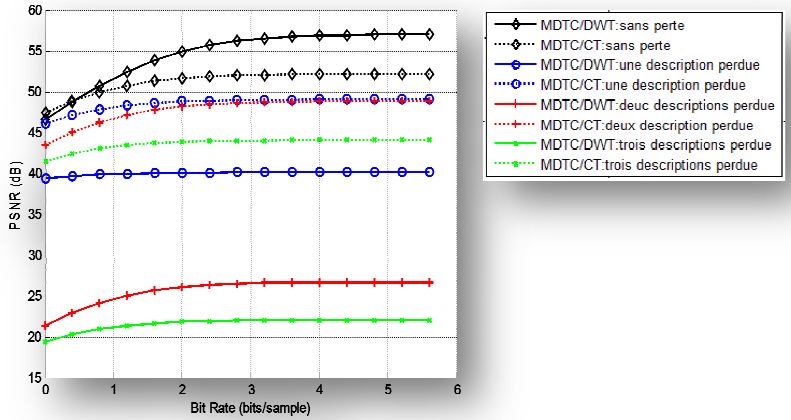
Nous présentons dans le tableau 4.2 le $PSNR$ calculé à un débit binaire $R = 2bps$ pour les deux codeurs *MDTC-DWT* et *MDTC-CT* dans différents cas de perte de descriptions utilisant plusieurs images en niveau de gris.

	PSNR (dB) pour MDTC/DWT					PSNR (dB) pour MDTC/CT				
Nombre de paquets perdus	Lena	Zelda	Peppers	Boat	Cameraman	Lena	Zelda	Peppers	Boat	Cameraman
0	55.00	61.13	51.44	51.30	57.60	52.18	58.85	49.54	45.14	50.85
1	40.15	55.11	40.14	41.86	47.16	49.09	58.20	44.23	43.18	47.16
2	26.70	47.25	36.36	33.63	44.23	48.87	56.01	43.99	42.74	46.48
3	22.08	24.89	20.00	22.70	39.27	44.11	51.59	42.34	38.53	43.99

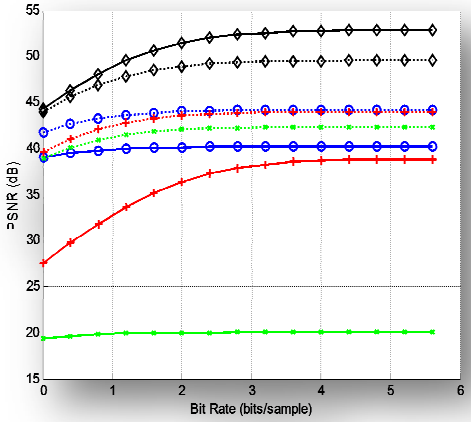
Tableau 4.2. $PSNR$(dB) en fonction du nombre de descriptions perdues pour *MDTC-DWT* et *MDTC-CT*.

D'après ces résultats les valeurs de $PSNR$ décroient rapidement en fonction de perte de descriptions pour le codeur *MDTC-DWT*, tandis que le codeur *MDTC-CT* donne des valeurs de $PSNR$ assez élevées dans les différents cas de perte de paquets. D'après le tableau, malgré la perte de plus de descriptions la *CT* engendre moins de dégradations et reste toujours meilleure que la *DWT* dans le cas de perte d'une seule ou plusieurs descriptions dans un schéma *MDTC*.

(a)



(b)



(c)

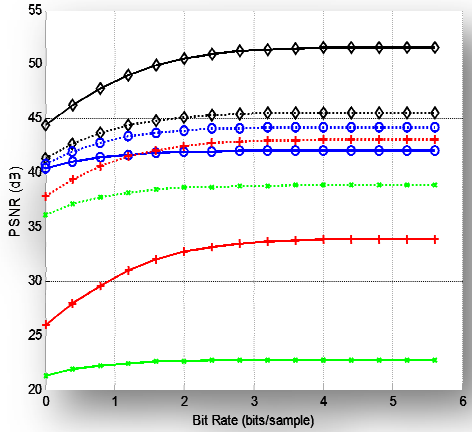


Figure 4.7. *PSNR* en fonction du débit binaire pour les images : (a) Lena (b) Peppers (c) Boat.

Les résultats menés sur les courbes précédentes présentent le $PSNR$(dB) latéral en fonction du débit R estimé à partir des entropies scalaires. Il est claire de ces courbes que la *CT* augmente les performances du codeur *MDTC* par rapport à *DWT* en fonction du nombre de descriptions perdues.

IV.2.5. Codage MDC basé sur la NSCT et l'approximation de poursuite adaptative orthogonale OMP :

Dans la *NSCT* toutes les opérations de sous-échantillonnage sont éliminées afin d'éviter le phénomène d'aliasing. Pour une image $k \times m$ avec une décomposition à 3 niveaux et 4 directions la *NSCT* offre une sous bande d'approximation et 3×4 sous bandes directionnelles. L'idée de la contribution de cette section repose sur une combinaison de la transformée redondante *NSCT* et l'algorithme d'approximation sparse *OMP* dans le but de réduire significativement le nombre de coefficients issus de la *NSCT* . L'*OMP* a été choisi pour la rapidité et facilité de son exécution, ce qui été démontré par Kaur et al [130].

Notre méthode proposée sur la figure 4.8, suit les mêmes étapes d'un codeur *MDCT-DWT* après avoir réduire le nombre de coefficients de la *NSCT* avec *OMP*.

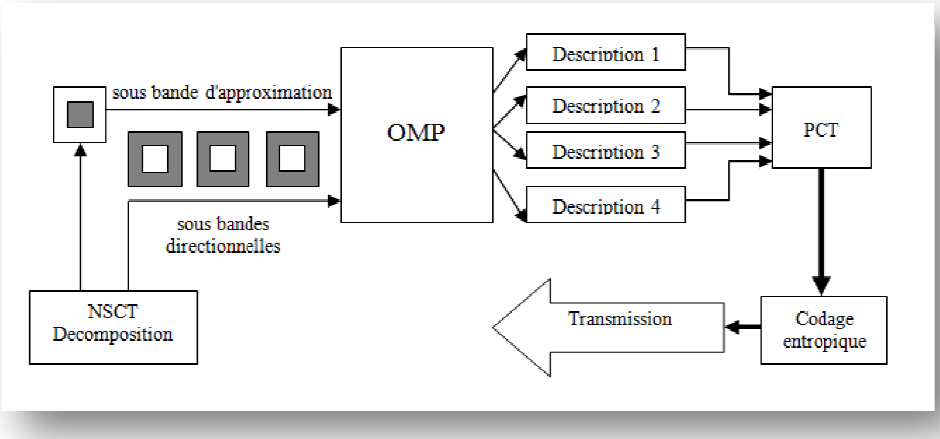


Figure 4.8. Approximation *OMP* des sous bandes *NSCT* pour le codeur *MDTC*.

Dans le tableau suivant, on commence par comparer l'algorithme d'approximation parcimonieuse *OMP* au *SPM* par le calcul du temps d'exécution et le taux de parcimonie SR en fonction du partitionnement en blocs, soit une image de taille $N \times N$ pixels transformée en une image parcimonieuse de K coefficients ($K < N$) le taux SR est donné par :

$$SR = \frac{Nombre\ de\ pixels}{Nombre\ de\ coefficients} = \frac{N^2}{K} \tag{4.5}$$

	OMP				SPM			
Taille de bloc	Boat		Lena		Boat		Lena	
	SR	T (sec)	SR	T (sec)	SR	T (sec)	SR	T (sec)
8	22.50	15.18	29.47	11.43	22.49	16.61	29.47	11.71
16	36.81	21.90	53.06	14.81	36.78	23.63	53.07	15.45
32	46.01	89.62	69.88	44.82	46.04	91.70	69.97	56.29

Tableau 4.3. Taux SR et temps d'exécution T en fonction de partitionnement en bloc pour *OMP* et *SPM*.

Le tableau 4.3 et la figure 4.9 illustrent les résultats de l'application du codeur *MDTC-NSCT-OMP* par rapport au *MDTC-DWT*. Nous observons de ces résultats l'efficacité de notre schéma en terme de $PSNR$(dB) qui est toujours élevé en cas de perte de descriptions.

Image	*No packet lost*		*One packet lost*		*Two packets lost*		*Three packets lost*	
	MDTC/ DWT	*MDTC/ NSCT*	*MDTC/ DWT*	*MDTC/ NSCT*	*MDTC/ DWT*	*MDTC/ NSCT*	*MDTC/ DWT*	*MDTC/ NSCT*
Lena	51.9044	59.9818	47.0825	55.5632	33.7629	45.9010	22.0660	34.0144
Boat	48.5996	54.9127	44.3451	51.3089	32.4978	43.4178	22.6637	34.5526
Cameraman	38.2272	54.2173	36.2595	51.1319	24.6921	39.8859	17.3969	29.4660
Mandril	47.7059	58.6980	45.0017	54.5192	34.3345	46.0407	25.6687	37.3537
Pirate	48.5735	55.9058	44.7491	52.4757	33.2477	44.4760	22.2382	34.1783
Livingroom	48.0290	55.2236	45.0797	52.1659	33.5962	44.2491	23.6095	35.4748

Tableau 4.4. $PSNR$(dB) obtenu par le codeur *MDTC-NSCT-OMP* et *MDTC-DWT*.

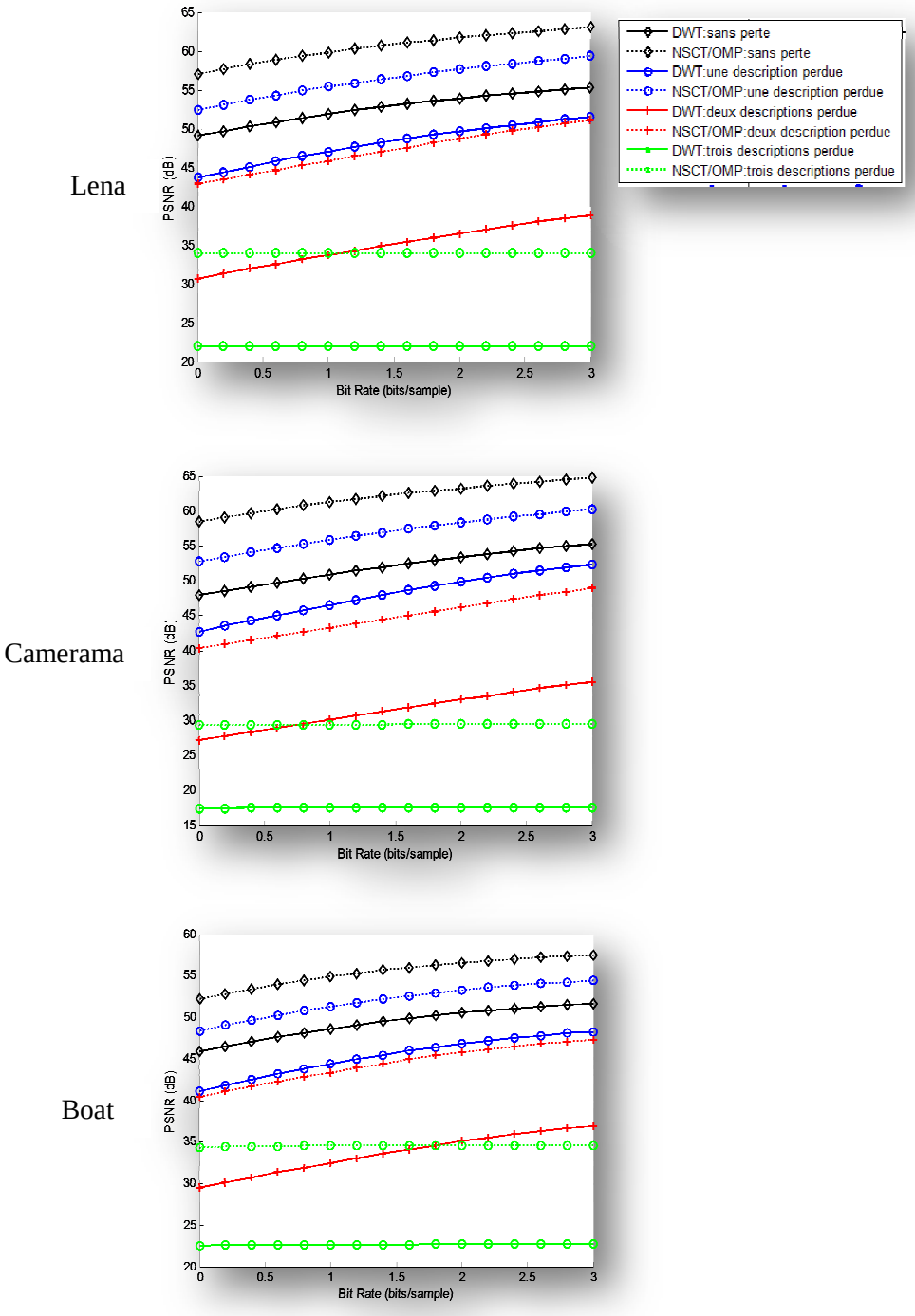


Figure 4.9. $PSNR(dB)$ en fonction du débit et le nombre de descriptions perdues.

IV.3. Deuxième Partie

Extension du codage par descriptions multiples à la vidéo

L'objectif de cette thèse est l'étude de codage *MDC* et plus particulièrement le codeur vidéo à descriptions multiples. Nous explorons un schéma de descriptions multiples basé sur une décomposition spatiale de la séquence vidéo utilisant la transformée redondante *NSCT*. Plusieurs travaux ont été consacrés récemment pour le codage vidéo *MD* et plus particulièrement basés sur des transformés redondantes. Wang et al et yang et al proposent d'introduire la transformé *DDWT* (Dual Tree Wavelet Transform) dans un schéma de codage des images et vidéo [131][132]. Le challenge de cette méthode est la redondance présente dans les trames issues de la décomposition *DDWT*, pour remédier à ce problème, les auteurs ont exploité l'algorithme d'approximation parcimonieuse *NS* (Noise Shaping) proposé par Kingsbury afin de réduire le nombre de coefficients pour représenter l'image.

dans cette partie on présente les résultats de simulations de nos méthodes proposées et basées sur la représentation redondante *NSCT*. Au sein de ce codeur nous veillerons particulièrement à :

- ✓ Augmenter l'efficacité du codeur *MDC* en exploitant une nouvelle transformée directionnelle et redondante.

- ✓ voir les performances et la robustesse de tel codeur basé sur un schéma de pré et post traitement.

IV.3.1. Codage vidéo à descriptions multiples par décalage de trame :

Nous proposons dans cette section, un algorithme de codage vidéo *MD* qui repose sur la transformée *NSCT* et l'algorithme *OMP* pour une approximation parcimonieuse de l'image du fait de la présence des coefficients redondants. Une transformation en contourlets non sous échantillonnée est appliquée aux images d'entrée, qui seront ensuite approximées avec l'*OMP*.

Ainsi, nous avons appliqué le principe de décalage dans l'ordre des trames de la séquence proposé par Lin et al pour la vidéo stéréoscopique [133]. Ce principe est bénéfique dans le cas de perte de quelques trames ou images à partir d'une description dans le but de la récupérer à partir de sa position décalée dans l'autre description.

Nous avons ensuite voulu comparer notre schéma *MDC-NSCT* et *MDC DWT* avec et sans décalage des sous bandes des descriptions à transmettre. Pour ce faire, on a tester une séquence *Foreman QCIF* de dix trames suivant les étapes du codage *MD* :

- ✓ Décomposer les images source avec *NSCT* en quatre sous bandes.
- ✓ Chaque sous bande est approximée avec l'*OMP*.
- ✓ Quantification uniforme des coefficients résultants.
- ✓ Codage entropique de ces vecteurs pour former enfin les descriptions à transmettre.

Le principe de décalage d'ordre des trames est illustré sur la figure 4.10.

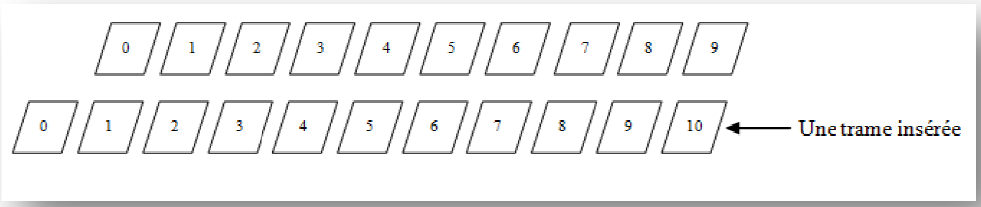
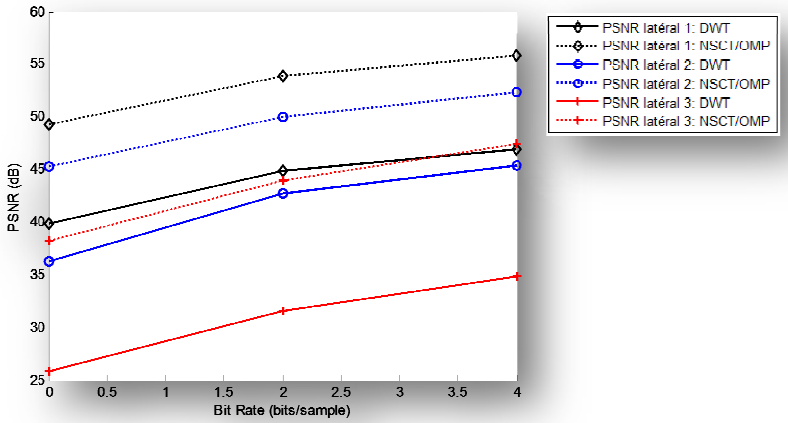


Figure 4.10. Schéma de décalage dans l'ordre de la séquence vidéo.

Dans notre cas de simulation, on a considéré l'exemple de perte de la 3è*me*, 4è*me*, 5è*me* et 6è*me* trame de la 1è*re* description, 2è*me* description, 3è*me* description et 4è*me* description respectivement. Les figures 4.11 présentent le tracé du *PSNR* en fonction du débit binaire *R* et le nombre de descriptions perdues.

(a)



(b)

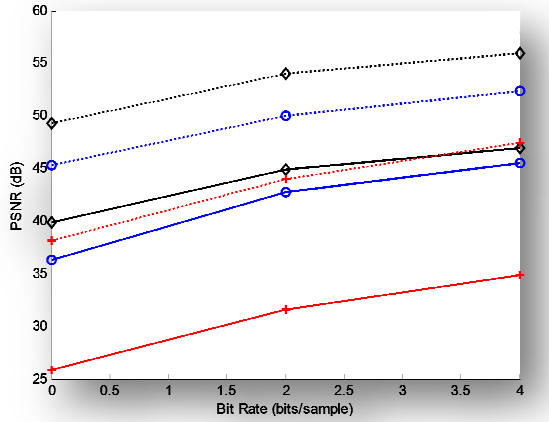


Figure 4.11. *PSNR* en fonction du débit pour *MDC-NSCT-OMP* et *MDC-DWT*. (a) sans décalage d'ordre de trames (b) avec décalage d'ordre de trames.

Les figures précédentes montrent que le $PSNR$ présente des valeurs importantes pour le *MDC-NSCT-OMP* que le *MDC-DWT* avec et sans considération du décalage dans la séquence. Nous avons enfin précisé sur le tableau 4.5 à un débit $R = 2bits/échantillons$.

	PSNR1	PSNR2	PSNR3	PSNR4
DWT	46.2786	44.5718	33.7145	18.0874
NSCT	55.2156	51.5389	46.1780	29.9439
NSCT avec décalage	55.2368	51.5410	46.1790	29.9415

Tableau 4.5. $PSNR$(dB) en fonction du nombre de descriptions perdues.

IV.3.2. Pré et Post traitement spatiale pour un schéma MDC : Application à la vidéo.

Bien que les techniques de codage vidéo à descriptions multiples mentionnées dans la section précédente apportent des résultats satisfaisants. Pour vaincre les limitations fournies par le *MDC*, le principe de pré et post traitement des images appliqué au codage *MD* [73] consiste à sur-échantillonner le signal d'entrée pour augmenter la corrélation entre les pixels adjacents, les images sur-échantillonnées sont ensuite décomposées pour générer des sous images qui peuvent être codées indépendamment. Dans un autre travail, Gallant et al [74] ont étendu leur approche de pré et post traitement spatiale à la vidéo, en sur-échantillonnant les images de la séquence pour introduire de la redondance à la séquence originale suivi d'une décomposition paire et impaire de la séquence sur-échantillonnée. Ces sous images peuvent être considérées comme deux séquences vidéos distinctes à coder indépendamment. Les séquences vidéo résultantes sont alors codées avec le standard de compression vidéo *H.263*. Dans [72], un codeur vidéo à descriptions multiples basé sur un sous-échantillonnage spatiale a été proposé, une redondance est ajoutée avec remplissage de zéros dans le domaine *DCT* pour augmenter la corrélation inter-pixels.

Un autre schéma a été proposé par Bernardini et al [134], les descriptions sont générées utilisant un banc de filtre redondant appliqué à chaque trame de la séquence d'entrée, deux filtres $h_0(n)$ et $h_1(n)$ divisent les lignes paires et impaires de la trame, un troisième filtre $h_2(n)$ a été appliqué afin de générer une troisième description. Ces descriptions sont alors codées avec le standard *H.264/AVC*.

Dans cet alinéa, on propose une approche *MD* basée sur un autre principe de pré et post traitement spatiale des images de la séquence vidéo d'entrée. Les descriptions sont générées par le banc de filtres pyramidale non sous-échantillonné *NSP* de la *NSCT* qui donne deux sous bandes de hautes et basses fréquences pour chaque trame originale, ces deux sous bandes sont ensuite séparées en trame paire et impaire pour générer deux descriptions.

Une première description est formée à partir des échantillons pairs de la sous bande basses fréquences et la sous bande hautes fréquences, tandis que la deuxième description est construite avec les échantillons impaires des sous bandes de basses et hautes fréquences. Les descriptions ainsi obtenues seront codées avec le standard *H.264/AVC*. Le codeur *MD* proposé est présenté sur la figure 4.12.

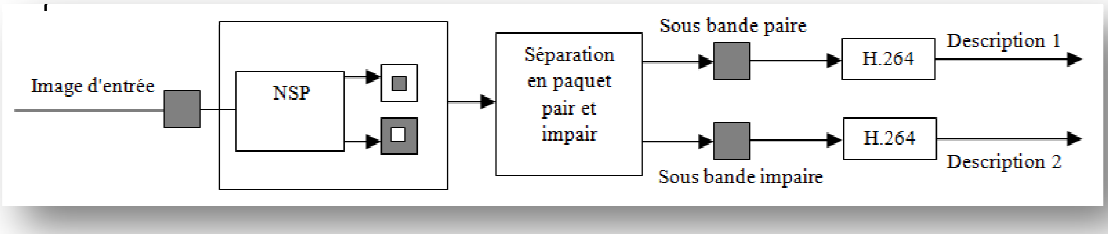


Figure 4.12. Diagramme en blocs du schéma *MD* proposé.

Dans cette section, on a présenté les résultats relatifs aux évaluations objectives de la qualité de reconstruction et de la robustesse du schéma proposé utilisant la séquence vidéo *Foreman QCIF* de 10 images (figure 4.13 et figure 4.14). La méthode proposée pour l'application du codage vidéo *MD* a été achevée avec le profil de base du standard *H.264*.

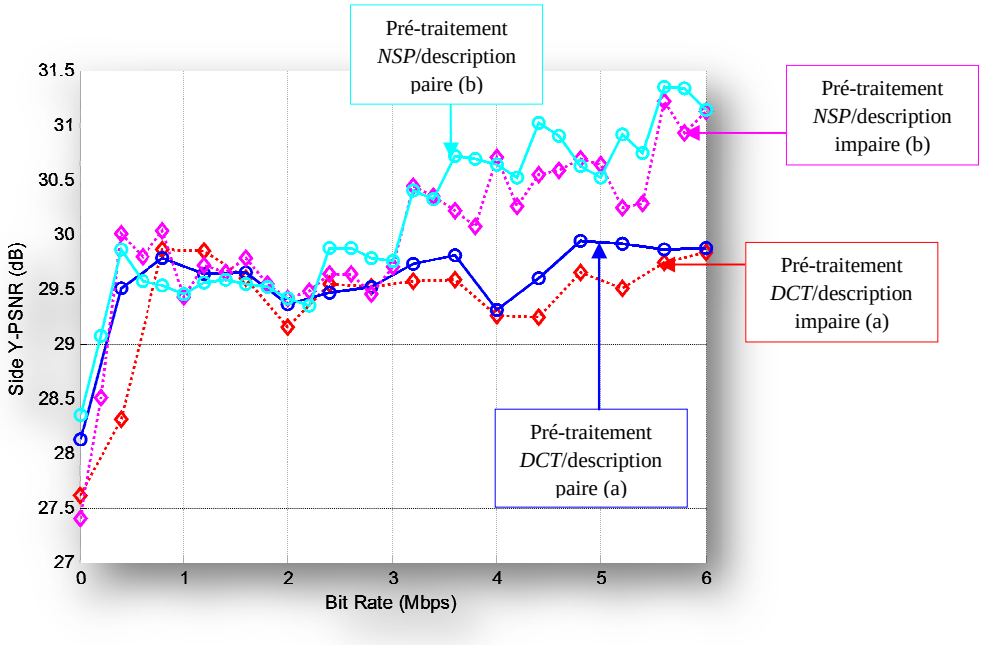


Figure 4.13. $Y-PSNR$ latéral en fonction du débit utilisant la séquence *Foreman QCIF* pour : (a) pré-traitement avec sur-échantillonnage spatiale dans le domaine *DCT* (b) pré-traitement avec *NSP*.

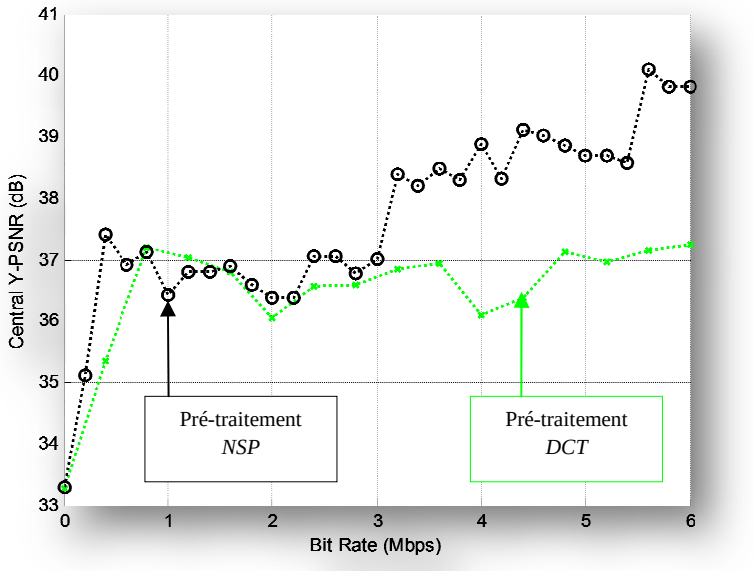


Figure 4.14. $Y-PSNR$ central en fonction du débit utilisant la séquence *Foreman QCIF* pour : (a) pré-traitement avec sur-échantillonnage spatiale dans le domaine *DCT* (b) pré -traitement avec *NSP*.

Les figures 4.13 et 4.14 illustrent une comparaison du $Y - PSNR$ latéral et central respectivement en fonction du débit pour la méthode de pré traitement basée sur un sur-échantillonnage dans le domaine *DCT* avec remplissage de zéros [74] et notre méthode proposée de pré-traitement qui repose sur la décomposition pyramidale *NSP* (premier étage de la transformée *NSCT*), considérant le scénario à deux descriptions. En effet, les valeurs de $PSNR$ obtenues par notre mesure montrent la possibilité de trouver un bon compromis entre l'efficacité de codage et la qualité de reconstruction quand une seule description est disponible. Il est claire de ces figures que notre schéma est plus performant de la méthode proposée dans [74]. Sur le tableau 4.6. on considère le $Y - PSNR(dB)$ central des trames reconstruites obtenues à partir de deux descriptions x et y en fonction du débit binaire. Un gain de $1.5dB$ à un débit de $5\,Mbs$ est obtenu avec le codage vidéo *MD* utilisant le banc de filtres pyramidale, ceci est principalement du au sur-échantillonnage présent dans la LP, cette propriété optimise l'efficacité de l'estimateur quand une seule description est reçue.

Méthode / ***R* (Mbit/échantillons)**	**Pré-traitement *NSP***	**Pré-traitement *DCT***
1	36.4282	37.1918
4	38.9737	36.1031

Tableau 4.6. Comparaison des valeurs de $Y - PSNR(dB)$ central à un débit de $1bps$ et $4bps$ pour les deux méthodes de pré et post traitement à base de *DCT* et *NSP*.

IV.4. Conclusion :

Dans ce chapitre, On a présenté aussi les contributions majeures de cette thèse. On a proposé un nouveau schéma *MDTC* basé sur une transformée redondante *NSCT* qui est effective pour la compression des images. On a fait une comparaison entre *NSCT-DWT* et *NSCT-DWT-OMP*, pour différentes images en niveau de gris où un gain en terme de $PSNR$ est obtenu avec notre proposition. Á partir des expérimentations, la *NSCT* a montré que une transformée redondante peut donner de meilleurs résultats et performances ainsi une qualité visuelle de l'image reconstruite dans des schémas de compression. Nous avons élaboré un examen de certaines solutions de pré et post traitement spatial de l'image permettant de remédier aux limitations du codage *MD* et améliorer ces performances. Nous avons aussi exhibé brièvement quelques solutions bibliographiques de recherche pouvant être de nouvelles perspectives. Il s'agit de nouvelle approche de codage *MD* basé sur le pré et post traitement temporel de la séquence d'entrée au codeur. Cette méthode vise à exploiter la redondance Inter-images de la séquence vidéo.

Dans cette thèse nous nous sommes intéressé au codage à descriptions multiples, des représentations multirésolustion de l'image et les algorithmes d'approximations qui fournissent une représentation parcimonieuse, ces algorithmes permettent d'analyser les données très redondantes à l'aide de dictionnaire d'atomes.

L'objet a été d'une part de réduire le nombre de coefficients de la transformée *NSCT* et l'exploitation de la corrélation entre les descriptions d'autre part. Cette stratégie avait pour but d'améliorer les performances du codage *MD*. À cette fin nous avons considéré des modèles basés sur l'*OMP* et une transformée de corrélation *PCT* pour introduire une certaine quantité de corrélation entre les descriptions. Notre algorithme *MDTC-NSCT-OMP* a été comparé au schéma *MDTC-DWT* et qui a prouvé des résultats significatifs par rapport à ce dernier. Les résultats de nos expérimentations présentent un gain considérable en terme de *PSNR* avec *CT* et *NSCT* qu'avec la décomposition en ondelettes utilisant différentes images en niveau de gris considérant une quantification uniforme et un codage entropique des descriptions à transmettre.

Nous avons étudié aussi un schéma *MDC* basé sur le pré traitement des images d'entrée au codeur en exploitant la caractéristique redondante dû à la présence de sur-échantillonnage dans le banc de filtres de la *NSCT*. Cette approche a été comparé à une méthode de pré et post traitement spatiale de la séquence vidéo qui vise à sur-échantillonner les images d'entrée et séparer les résultantes en descriptions paire et impaire.

Depuis l’apparition du multimédia, les méthodes de compression vidéo n’ont cessé de s’améliorer, afin d’optimiser le stockage et la diffusion à travers les réseaux de séquences audiovisuelles. On l’a vu, la norme de codage vidéo *H.264* est le fruit des recherches développées conjointement par Video Coding Experts Group (*VCEG*) ainsi que Moving Picture Experts Group (*MPEG*). Cette technologie employée est aussi connue sous le nom Advanced Vidéo Coding (*AVC*).

Nous avons introduit les principes générales du codeur *H.264* dans le deuxième chapitre. Le *H.264* a été développé pour compenser la grande quantité de données à transmettre, ainsi qu'améliorer le taux de compression en comparaison avec les anciens standards.

Le standard *H.264* permet de réduire la quantité de données à 50% tout en conservant presque la même qualité vidéo. L'encodeur *H.264* est un hybride de deux types de codage en Intra, pour réduire les redondances spatiales au niveau de l'image elle même et codage en Inter pour éliminer les redondances temporelles entre les images successives.

La comparaison de nos approches a bien montré sa performance pour la transmission des images et vidéo à descriptions multiples.

D'un point de vu perspectives, Il s'agit de nouvelle approche de codage *MD* basée sur le pré et post traitement temporel de la séquence d'entrée au codeur, notons que avec la migration vers la résolution Haute Définition (*HD*) les améliorations au niveau du *H.264* ont engendré une grande complexité de calcul, ce qui rend plus difficile le codage de la vidéo à haute définition. Cette méthode vise à exploiter la redondance Inter-images de la séquence et les descriptions seront codées avec le nouveau standard High Efficiency Video Coding (*H.265*/*HEVC*), développé conjointement par les groupes Video Coding Experts Group (*VCEG*) et Moving Picture Experts Group (*MPEG*) de codage vidéo avec une large gamme de résolutions à partir de basse résolution à la haute définition ($4K$ ou $8K$), ce standard étant capable de réduire le débit binaire à 50% , en conservant la même qualité de la vidéo.

[1] A. Naimi and K. Belloulata. "Multiple Description Image Coding using Contourlet Transform"Image Compression based Multiple Description Transform Coding using NSCT and OMP Approximation". In press for publication in International Journal of Computational Vision and Robotics (IJCVR).

[2] A. Naimi and K. Belloulata. "Multiple Description Image Coding using Contourlet Transform". IEEE International Conference on Electrical Engineering, Boumerdes, Algeria, December 2015, In press for publication in International Journal of Computer Aided Engineering and Technology (IJCAET).

[3] A. Naimi and K. Belloulata. "Multiple Description Image Coding Comparison for Multiresolution Image Representation". Proceedings International Conference on Digital Information Processing, Data Mining, and Wireless Communications, Dubai 2015.

[4] A. Naimi and K. Belloulata. "Multiple Description Video Coding based on NSCT and Staggered Frame Order". International Conference on Advanced Communication Systems and Signal Processing, pp. 670-675, November 2015.

[1] T. Tanab and N. Farvardin. "Subband Image Coding using Entropy Coded Quantization over Noisy Channels". IEEE Journal Selected Areas on Communication, vol. 10, pp. 926-943, June 1992.

[2] I. Daubechies. "Ten Lectures on Wavelets". SIAM, Philadelphia, PA, 1992.

[3] C. Heil and D.F. Walnut. "Fundamental Papers in Wavelet Theory" pp. 912, 2006.

[4] J. Radon. "On the Determination of Functions from their Integral Values along Certain Manifolds". In Reports Saxon Academy of Sciences, Leipzig Math Nat, vol. 69, pp. 262-277, 1917.

[5] F. Matus and J. Flusser. "Image Representation via a Finite Radon Transform". IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 15, no. 10, pp. 996-1006, 1993.

[6] E.J. Candes. "Ridgelets : Theory and Applications". Thèse de Doctorat, Départment de statistiques, Université de Stanford, 1998.

[7] M.N. Do and M. Vetterli. "Orthonormal Finite Ridgelet Transform for Image Compression". IEEE International Conference on Image Processing, vol. 2, pp. 367-370, 2000.

[8] D.L. Donoho. "Orthonormal Ridgelets and Linear Singularities". Tech Rep, Stanford University, 1998.

[9] E.J. Candes. "Ridgelets: The Key to Higher Dimensional Intermittency?". Philosophical Transactions of the Royal Society, vol. 357, pp. 2495-2509, 1999.

[10] E.J. Candes. "Monoscale Ridgelets for The Representation of Images with Edges". Tech Rep, Department of Statistics, Stanford University, 1999.

[11] D.L. Donoho. "Orthonormal Ridgelets and Linear Singularities". SIAM, Journal of Mathematical Analysis, vol. 31, no. 5, pp. 1062-1099, 2000.

[12] M.N. Do and M. Vetterli. "The Finite Ridgelet Transform for Image Representation". IEEE Transactions on Image Processing, vol. 12, no. 1, pp. 16-28, 2003.

[13] E.J. Candes and D.L. Donoho. "Curvelets : A Surprisingly Effective Non-Adaptive Representation for Objects with Edges", International Conference on Curves and Surfaces, Vol. 2, pp. 1-10, Saint Malo, France, 1999.

[14] E.J. Candes and D.L. Donoho. "New Tight Frames of Curvelets and Optimal Representations of Objects with Smooth Singularities", Tech Rep, Stanford University, 2002.

[15] J.L. Starck, E. Candes and D L. Donoho. "The Curvelet Transform for Image Denoising". IEEE Transactions on Image Processing, vol. 11, no. 6, pp. 670-684, 2002.

[16] D.L. Donoho and X. Huo. "Beamlet Pyramids : A New Form of Multiresolution Analysis Suited for Extracting Lines, Curves and Objects From Very Noisy Image Data". In SPIE Conference on Wavelet Applications in Signal and Image Processing, vol. 4119, pp. 434-444, 2000.

[17] D.L. Donoho. "Wedgelets : Nearly Minimax Estimation of Edges". In Annals of Statistics, vol. 27, no. 3, pp. 859-897, 1999.

[18] M. Wakin, J. Romberg, H. Choi, and R. Baraniuk. "Rate-Distortion Optimized Image Compression using Wedgelets". In IEEE International Conference on Image Processing,vol. 3, Rochester, NY, United States, Sept 2002.

[19] Z. Xiong, K. Ramchandran and M.T. Orchard."Space-Frequency Quantization for Wavelet Image Coding". IEEE Transactions on Image Processing, vol. 6, no. 5, pp. 677-693, 1997.

[20] V. Chandrasekaran, M. Wakin, D. Baron and R. Baraniuk. "Surflets : A Sparse Representation for Multidimensional Functions Containing Smooth Discontinuities". IEEE International Symposium on Information Theory, Chicago, June 2004.

[21] E. Pennec and S. Mallat. "Image Compression with Geometrical Wavelets". IEEE International Conference on Image Processing, vol. 1, pp. 661-664, 2000.

[22] E. Pennec and S. Mallat. "Sparse Geometric Image Representation with Bandelets". IEEE Transactions on Image Processing, vol. 14, no. 4, pp. 423-438, 2003.

[23] C. Bernard and E. Le Pennec. "Adaptation of Regular Grid Filtering to Irregular Grids". Tech Rep, École polytechnique, Centre de Mathématiques Appliquées, Palaiseau Cedex, France, 2002.

[24] G. Peyré and S. Mallat. "Surface Compression with Geometric Bandelets". ACM Transactions on Graphics, Proceedings of ACM SIGGRAPH 2005, vol. 24, no. 3, pp.601-608, Aug 2005.

[25] M. Vetterli, V. Velisavljevic, B. Beferull and P.L. Dragotti. "Directionlets : Anisotropic Multi-directional Representation with Separable Filtering". IEEE Transactions on Image Processing, vol. 15, no. 7, pp. 1916-1933, 2004.

[26] V. Velisavljevic, B. Beferull, M. Vetterli, and P.L. Dragotti. "Approximation Power of Directionlets". IEEE International Conference on Image Processing, Sept 2005.

[27] S. Mallat. "Multiresolution Approximations and Wavelet Orthonormal Bases of L2(R)". Transactions of the American mathematical society, vol. 315, no. 1, pp. 69-87, 1989.

[28] M. Shapiro. "Embedded Image Coding using Zerotres of Wavelet Coefficients". IEEE Transactions on Signal Processing, vol. 41, pp. 3445-3462, Dec. 1993.

[29] A. Cohen. "Ondelettes et Traitement Numérique du Signal". Masson, pp. 205, Paris, 1992.

[30] A. Skodras, C. Christopoulos and T. Ebrahimi. "The JPEG2000 Still Image Compression Standard". IEEE Signal Processing Magazine, vol. 18, no. 5, pp. 36-58, Sep 2001.

[31] J. Daugman. "Two Dimensional Spectral Analysis of Cortical Receptive Field Profile". Vision Research, vol. 20, no. 10, pp. 847-856, 1980.

[32] A.B. Watson."The Cortex Transform : Rapid Computation of Simulated Neural Images". Computer Vision Graphics and Image Processing, vol. 39, no. 3, pp. 311-327, 1987.

[33] E.P. Simoncelli, W.T. Freeman, E.H. Adelson and D.J. Heeger. "Shiftable Multiscale Transforms". IEEE Transactions on Information Theory, Special Issue on Wavelet Transforms and Multiresolution Signal Analysis, vol. 38, no. 2, pp. 587-607, 1992.

[34] J.P. Antoine, P. Carrette, R. Murenzi and B. Piette. "Image Analysis with Two Dimensional Continuous Wavelet Transform". Signal Processing, vol. 31, no. 3, pp. 241-272, 1993.

[35] F.G. Meyer and R.R. Coifman. "Brushlets : A Tool for Directional Image Analysis and Image Compression". Journal of Applied and Computational Harmonic Analysis, vol. 5, no. 2, pp. 147-187, 1997.

[36] N. Kingsbury. "Complex Wavelets for Shift Invariant Analysis and Filtering of Signals". Journal of Applied and Computational Harmonic Analysis, vol. 10, no. 3, pp. 234-253, 2001.

[37] P.J. Burt and E.H. Adelson. "The Laplacian Pyramid as A Compact Image Code". IEEE Transactions on Communications, vol. 31, no. 4, pp. 532-540, 1983

[38] M. Vetterli and J. Kovacevic. "Wavelets and Subband Coding". Prentice Hall, pp. 522, 1995.

[39] M.N. Do and M. Vetterli. "Framing Pyramids". IEEE Transactions on Signal Processing, vol. 51, no. 9, pp. 2329-2342, 2003.

[40] R.H. Bamberger and M.J.T. Smith. "A Filter Bank for The Directional Decomposition of Images : Theory and Design". IEEE Transactions on Signal Processing, vol. 40, no. 4, pp. 882-893, April 1992.

[41] M. Vetterli. "Multidimensional Subband Coding : Some Theory and Algorithms" Signal Processing, vol. 6, no. 2, pp. 97-112, 1984.

[42] M.N. Do. "Directional Multiresolution Image Representations". Ph.D. dissertation, Swiss Federal Institute of Technology, Lausanne, Switzerland, December 2001.

[43] M.N. Do and M. Vetterli. "The Contourlet Transform : An Efficient Directional Multiresolution Image Representation". IEEE Transactions on Image Processing, vol. 14, no. 12, pp. 2091-2106, 2005.

[44] M.N. Do and M. Vetterli. "Pyramidal Directional Filter Banks and Curvelets". IEEE International Conference on Image Processing, Thessaloniki, Greece, Oct 2001.

[45] M.N. Do and M. Vetterli. "Contourlets : A Directional Multiresolution Image Representation". IEEE International Conference on Image Processing, Rochester, USA Sept 2002.

[46] A.L. Cunha, J. Zhou and M.N. Do. "The Nonsubsampled Contourlet Transform : Theory Design and Applications". IEEE Transactions on Image Processing, vol. 15, no. 10, pp. 3089-3101, 2006.

[47] M. BENABDELLAH. " Outils de Compression et de Crypto-compression : Applications aux images fixes et vidéo". Thèse de Doctorat, Faculté des Sciences, Rabat-Maroc, 2007.

[48] S.S. Pradhan, R. Puri and K. Ramchandran. "(n, k) Source Channel Erasure Codes : Can Parity Bits also Refine Quality ?". Conference on Information Sciences and Systems, March 2001.

[49] R. Puri, S.S. Pradhan and K. Ramchandran." n-Channel Symmetric Multiple Descriptions part ii : an Achievable Rate Distortion Region". IEEE Transactions on Information Theory, vol. 51, no. 4, pp. 1377-1392, 2005. Theory, vol. 51, pp. 1377-1392, April 2005.

[50] S.S. Pradhan and K. Ramchandran. "Distributed Source Coding using Syndromes (discus) : Design and Construction". Data Compression Conference, pp. 158-167, March 1999.

[51] F. Aurenhammer. "Voronoi Diagrams-A survey of A Fundamental Geometric Data Structure". ACM Comuting Surveys, vol. 23, n. 3, pp. 345-405, 1991.

[52] J.M. Pascal. "La Compression des Données ou Le Flux Multimédia Sous Contrainte ". Découverte N.0283, Décembre 2000.

[53] M. Irani, P. Anandan, J. Bergen, R. Kumar and S. Hsu. "Efficient Representations of Video Sequences and Their Applications". Signal Processing : Image Communication, vol. 8, no. 4, :327-351, 1996.

[54] Monique COLINET, " Multimédia : Images, Sons et Vidéos ". Université Notre Dame de Paix, CEFIS-FUNDP, Février 2001.

[55] F. M. Reza. "An Introduction to Information Theory ". Dover Publications, Inc, New-York, pp. 496, 1994.

[56] R.E. Blahut. "Digital Transmission of Information". Addison Wesley, pp. 562, 1990.

[57] P. Fiche, V. Ricordel and C. Labit. "Etude d'Algorithmes de Quantification Vectorielle Arborescente Pour La Compression d'Images Fixes". Rapport de recherche INRIA, 1994.

[58] G. Patané and M. Russo. "The Enhanced LBG Algorithm". Neural Networks vol. 14, no. 9, 1219-1237, 2001.

[59] C. Delgorge. "Proposition et Evaluation de Techniques de Compression d'Images Ultrasonores Dans Le Cadre d'Une Télé-Echographie Robotisée". Ph.D thesis, Université d'Orléans, 2005.

[60] O. Rioul. "Ondelettes Régulières : Application à la Compression d'Images Fixes". Ph.D thesis, Ecole Nationale Supérieure des Télécommunications, 1993.

[61] S.G. Mallat. "A Theory for Multiresolution Signal Decomposition : The Wavelet Representation". IEEE Transactions of Pattern Analysis and Machine Intelligence, vol. 2, no. 7, pp. 674-693, 1989.

[62] A.P. Bradley and F.W.M. Stentiford. "JPEG2000 and Region of Interest Coding". Digital Image Computing Techniques and Applications, Melbourne, Australia, January 2002.

[63] M. Banckaert, P. Gasser and M. Lavacry. "CD-I Digital Video et VHS : essais comparatifs", CNDP, 1994.

[64] M. Abdat. "Etudes des Techniques de Compression des Images Fixes et Amélioration de La Résistance aux Erreurs de Transmission". Ph.D thesis, Blida, Algerie, 1995.

[65] T. Huffman, D.Muller."Techniques de Compression des Signaux Vidéo Pour La Communication Multimédia". Revue des Télécommunications, vol. 4, pp. 402-410 1993.

[66] ITU-T Recommendation *H.264*. "Advanced Video Coding for Generic Audiovisual Services". International Telecommunication Union Standardization Sector, Novembre 2007.

[67] G. Robert. "Représentation et Codage de Séquences Vidéo Par Hybridation de Fractales et d'Éléments Finis". Thèse de l'université joseph Fourier, 1997.

[68] Y.L. Steve, C.Y. Kao, H.C. Kuo, J.W. Chen. "VLSI Design for Video Coding : *H.264/AVC* Encoding From Standard Specification to Chip". pp. 176, 2010.

[69] S.D. Servrto, K. Ramchandran, V.A. Vaishampayan and K. Nahrstedt. "Multiple Description Wavelet Based Image Coding". IEEE Transactions on Image Processing, vol. 9, no. 5, pp.813-826, 2000.

[70] H. Bai, Y. Zhao, C. Zhu. "Multiple Description Shifted Lattice Vector Quantization for Progressive Wavelet Image Coding". IEEE International Conference on Image Processing, Atlanta, pp. 797-800, 2006.

[71] E. El Gamal, T. Cover. "Achievable Rates for Multiple Descriptions". IEEE Transactions on Information Theory. vol. 28, no. 6, pp. 851-857, 1982.

[72] D. Wang, N. Canagarajah, D. Redmill and D. Bull. "Multiple Description Video Coding Based on Zero Padding". International Symposium on Circuits and Systems, vol. 2, pp. 205-208, May 2004.

[73] S. Shirani, M. Gallant and F. Kossentini. "Multiple Description Image Coding using Pre and Post Processing". IEEE International Conference on Information Technology: Coding and Computing, pp. 35-39, Las Vegas, April 2001.

[74] M. Gallant, S. Shirani and F. Kossentini. "Standard Compliant Multiple Description Video Coding". International Conference on Image processing, vol. 1, pp. 946-949, Thessaloniki 2001.

[75] J. Wolf, A. Wyner and J. Ziv. "Source Coding for Multiple Descriptions". Bell Labs Technical Journal, vol. 59, no. 8, pp. 1417-1426, 1980.

[76] L. Ozarow. "On A Source Coding Problem With Two Channels and Three Receivers". Bell Labs Technical Journal , vol. 59, no. 10, pp. 1909-1921, 1980.

[77] H. Witsenhausen. "On Source Networks With Minimal Breakdown Degradation". Bell Labs Technical Journal , vol. 59, no. 6, pp. 1083-1087, 1980.

[78] H. Witsenhausen. "Minimizing The Worst Case Distortion In Channel Splitting". Bell Labs Technical Journal, vol. 60, no. 8, pp. 1979-1983, 1981.

[79] Z. Zhang and T. Berger. "New Results In Binary Multiple Descriptions". IEEE Transactions on Information Theory, vol. 33, no. 4, pp. 502-521, 1987.

[80] R. Venkataramani, G.Kramer, and V. K. Goyal, "Multiple description coding with many channels". IEEE Transactions on Information Theory, vol. 49, no. 9, pp. 2016-2114, 2003.

[81] R. Puri, S.S. Pradhan and K. Ramchandran. "n-Channel Symmetric Multiple Descriptions : New Rate Regions". IEEE International Symposium on Information Theory, Palais de Beaulieu. Lausanne, Switzerland, 2002.

[82] R. Puri, S.S. Pradhan and K. Ramchandran. "n-Channel Multiple Descriptions : Theory and Constructions". Data Compression Conference, pp. 262-271, April 2002.

[83] A. Albanese, J. Blomer, J. Edmonds, M. Luby, and M. Sudan. "Priority Encoding Transmission". IEEE Transactions on Information Theory". vol. 42, no. 6, pp. 1737-1744, 1996.

[84] A. Wyner and J. Ziv. "The Rate Distortion Function for Source Coding With Side Information at The Decoder". IEEE Transactions on Information Theory, vol. 22, no. 1, pp. 1-10, 1976.

[85] V.A. Vaishampayan. "Design of Multiple Description Scalar Quantizers". IEEE Transactions on Information Theory". vol. 39, no. 3, pp. 821-834, 1993.

[86] T.Y. Berger and E.M. Reingold. "Index Assignment for Multichannel Communication Under Failure". IEEE Transactions on Information Theory". vol. 48, no. 10, pp. 2656-2668, 2002.

[87] V.A. Vaishampayan and J. Domaszewicz. "Design of Entropy Constrained Multiple Description Scalar Quantizers". IEEE Transactions on Information Theory, vol. 40, no. 1, pp. 245–250, 1994.

[88] L. Buzi, V.A. Vaishampayan and R. Laroia. "Design and Asymptotic Performance of A Structured Multiple Description Vector Quantizer". IEEE International Symposium on Information Theory, 1994.

[89] S.D. Servetto, V.A. Vaishampayan and N.J.A. Sloane. "Multiple Description Lattice Vector Quantization". Data Compression Conference, pp. 13-22, 1999.

[90] V.A. Vaishampayan, N.J.A. Sloane and S.D. Servetto. "Multiple Description Vector Quantization With Lattice Codebooks : Design and Analysis". IEEE Trans. Inform. Theory, vol. 47, no. 5, pp. 1718-1734, 2001.

[91] H.S. Malvar and D.H. Staelin. "The *LOT* : Transform Coding Without Blocking Effects". IEEE Transactions on Acoustics, Speech and Signal Processing, vol. 37, no. 4, pp. 553-559, 1989.

[92] D.M. Chung and Y. Wang. "Multiple Description Image Coding using Signal Decomposition and Reconstruction Based On Lapped Orthogonal Transforms". IEEE Transactions on Circuits and Systems for Video Technology, vol. 9, no. 6, pp. 895-908, 1999.

[93] J. Lebrun and M. Vetterli. "Balanced Multiwavelets Theory and Design". IEEE Transactions on Signal Processing, vol. 46, no. 4, pp. 1119-1125, 1998.

[94] J. Lebrun and M. Vetterli. "High-Order Balanced Multiwavelets : Theory, Factorization and Design". IEEE Transactions on Signal Processing, vol. 49, no. 9, pp. 1918-1930, 2001.

[95] Z. Chen, C. Guillemot and R. Ansari. "Multiple Description Coding using Multiwavelet Transforms". Picture Coding Symposium, 2001.

[96] Y. Wang, M.T. Orchard and A.R. Reibman. "Multiple Description Image Coding for Noisy Channels by Pairing Transform Coefficients". IEEE First Workshop on Multimedia Signal Processing, pp. 419-424, 1997.

[97] M.T. Orchard, Y. Wang, V. Vaishampayan and A.R. Reibman. "Redundancy Rate-Distortion Analysis of Multiple Description Coding using Pairwise Correlating Transforms". International Conference on Image Processing, vol. 1, pp. 608-611, Oct 1997.

[98] V.K. Goyal and J. Kovacevic. "Optimal Multiple Description Transform Coding of Gaussian Vectors". Data Compression Conference, pp. 388-397, March 1998.

[99] V.K. Goyal, M. Vetterli and J. Kovacevic. "Multiple Description Transform Coding : Robustness to Erasures using Tight Frame Expansions". IEEE International Symposium on Information Theory, Aug 1998.

[100] P.L. Dragotti, S.D. Servetto and M. Vetterli. "Analysis of Optimal Filter Banks for Multiple Description Coding". Data Compression Conference, pp. 323-332, March 2000.

[101] X. Yang and K. Ramchandran. "Optimal Multiple Description Subband Coding". International Conference on Image Processing, vol. 1, pp. 654-658, Oct 1998.

[102] G.M. Davis and J.M. Danskin. "Joint Source and Channel Coding for Image Transmission over Lossy Packet Networks". Applications of Digital Image Processing XIX, SPIE, pp. 376–387, 1996.

[103] Y. Wang, A.R. Reibman, and S. Lin. "Multiple Description Coding for Video Delivery". IEEE, vol. 93, no. 1, pp. 57-70, 2005.

[104] S. McCanne and M. Vetterli. "Joint Source/Channel Coding for Multicast Packet Video". International Conference on Image Processing, vol. 1, pp. 25-28, Oct 1995.

[105] S. McCanne, M. Vetterli and V. Jacobson. "Low-Complexity Video Coding for Receiverdriven Layered Multicast. "IEEE Journal on Selected Areas in Communications, vol. 15, no. 6, pp. 983-1001, 1997.

[106] J.G. Apostolopoulos. "Reliable Video Communication over Lossy Packet Networks using Multiple State Encoding and Path Diversity". Visual Communications and Image Processing, vol. 4310, no. 1, pp. 392-409, 2000.

[107] M. Pereira, M. Antonini and M. Barlaud. "Multiple Description Image and Video Coding for Wireless Channels". Signal Processing : Image Communication, Special Issue on Recent Advances in Wireless Video, vol. 18, no. 10, pp. 925-945, 2003.

[108] M.U. Pereira, M. Antonini and M. Barlaud. "Channel Adapted Multiple Description Coding Scheme using Wavelet Transform". International Conference on Image Processing, Sept. 2002.

[109] C. Parisot, M. Antonini and M. Barlaud. "3d Scan Based Wavelet Transform for Video Coding". IEEE Fourth Workshop on Multimedia Signal Processing, pp. 403-408, Oct 2001.

[110] ITU Telecom Standardization Sector of ITU. "Video Coding for Low Bitrate Communication". ITU-T Recommendation *H.263* Version 2, January 1998.

[111] H. Bai, Y. Zhao and C. Zhu. "Low Complexity Pre-Post-Processing Multiple Description Coding for Video Streaming". IEEE International Conference on Multimedia and Expro, pp. 1331-1334, Beijing, July 2007.

[112] J.A. Trop. "Topics in Sparse Approximation". Ph.D thesis, University of Texas at Austin, August 2004.

[113] R.M.F.i Ventura. "Sparse Image Approximation with Application To Flexible Image Coding". Ph.D thesis, EPFL, September 2005.

[114] R. Gribonval. "Sur quelques problèmes mathématiques de modélisation parcimonieuse". PhD thesis, Université de Rennes 1, October 2007.

[115] S. Mallat and Z. Zhang. "Matching Pursuits with Time-Frequency Dictionaries". IEEE Transactions on Signal Processing, vol. 41, no. 12, pp. 3397-3415, 1993.

[116] F. Bergeaud and S. Mallat. "Matching Pursuit of Images". International Conference of Image Processing, pp. 53-56, 1995.

[117] S.S. Chen, D.L. Donoho and Michael A. Saunders. "Atomic Decomposition by Basis Pursuit". SIAM Journal on Scientific Computing, vol. 20, no. 1, pp. 33–61, 1999.

[118] S.S. Chen. "Basis Pursuit". Ph.D thesis, Stanford University, Nov 1995.

[119] D. Needell. "Topics in Compressed Sensing". Ph.D thesis, University of California, 2009.

[120] D.L. Donoho. "Unconditional Bases are Optimal Bases for Data Compression and For Statistical Estimation". Applied and Computational Harmonic Analysis, vol. 1, no. 1, pp. 100-115, 1993.

[121] V. Vaishampayan. "Design of Multiple Description Scalar Quantizers". IEEE Transactions on Information Theory, vol. 39, no. 3, pp. 821–834, 1993.

[122] Y. Wang, M. Orchard and A. Reibman. "Multiple Description Coding using Pairwise Correlating Transforms". IEEE Transactions on Image Processing, vol. 10, no. 3, pp. 351–366, 2001.

[123] Y. Wang, M. Orchard and A. Reibman. "Optimal Pairwise Correlating Transforms for Multiple Description Coding". IEEE International Conference on Image Processing, 1998.

[124] V.K. Goyal, J. Kovacevic. "Generalized Multiple Description Coding with Correlating Transforms". IEEE Transactions on Information Theory, vol. 47, no. 6, pp. 2199–2224, 2011.

[125] S. Sumohana, J.L. Channappayya,W.H.J. Robert and A.C. Bovik. "Frame Based Multiple Description Image Coding in The Wavelet Domain". International Conference on Image Processing, 2005.

[126] K. Khelil, R.E. Bekka, A. Djebbari and J.M. Rouvaen. "Multiple Description Wavelet Based Image Coding using Correlating Transforms". International Journal of Electronics and Communications, vol. 61, no. 6, pp. 411-417, 2007.

[127] V.K. Goyal, J. Kovacevic, R. Arean and M. Vetterli. "Multiple Description Transform Coding of Images". International Conference on Image Processing, pp. 674-678, 1998.

[128] A. Naimi and K. Belloulata. "Multiple Description Transform Coding Comparison for Multiresolution Image Representation". International Conference on Digital Information Processing, Data Mining and Wireless Communications". pp. 93-99, Dubai, 2015.

[129] A. Naimi and K. Belloulata. "Multiple Description Image Coding using Contourlet Transform". IEEE International Conference on Electrical Engineering, Boumerdes, Algeria, December 2015, In press for publication in International Journal of Computer Aided Engineering and Technology (IJCAET).

[130] A. Kaur and S. Budhiraja. "Wavelet Based Sparse Image Recovery Via Orthogonal Matching Pursuit". IEEE International Conference on Recent Advances in Engineering and Computational Sciences, pp. 1-5, March 2014.

[131] B. Wang,Y. Wang, I. Selesnick and A. Vetro. "Video Coding Using 3D Dual-Tree Wavelet Transform". EURASIP Journal on Image and Video Processing, Vol. 2007, pp. 15.

[132] J. Yang, Y. Wang, W. Xu and Qionghai Dai. "Image Coding Using Dual-Tree Discrete Wavelet Transform". IEEE transaction on Image Processing, vol. 17, no. 9, pp. 1555-1569, 2008.

[133] C. Lin, Y. Zhao, T. Tillo and J. Xiao. "Multiple Description Coding for Stereoscopic Videos with Stagger Frame Order". IEEE Transactions on Circuits and Systems for Video Technology, Vol. 25, no. 6, pp. 1016-1025, 2015.

[134] B. Bernardini, R. Rinaldo, A. Tonello and A. Vitali. "Frame based Multiple Description for Multimedia Transmission over Wireless Networks". International Syposium on Wireless Personal Miltimedia Communications, vol. 2, pp. 529-533, Italy, September 2004.

[135] M. HRARTI. " Optimisation du Controle de débit de H264/AVC basé sur une nouvelle Modelisation Débit-Quantification et Une Allocation Selective de Débit ", Thèse de Doctorat, Facultée des Sciences de Rabat et Facultée des Sciences Fondamentales et Appliquées de Poitiers, 2011.

[136] O. CRAVE. " Approches Théoriques en Codage Vidéo Robuste Multi-terminal ". Thèse de Doctorat, Télécoms ParisTech, 2008.