

N° d'ordre :

REPUBLIQUE ALGERIENNE DEMOCRATIQUE & POPULAIRE
MINISTRE DE L'ENSEIGNEMENT SUPERIEUR & DE LA RECHERCHE
SCIENTIFIQUE



UNIVERSITE DJILLALI LIABES
FACULTE DES SCIENCES EXACTES
SIDI BEL ABBÈS

THESE DE DOCTORAT

Présentée par

OUDDANE Samira

Spécialité : Informatique

Option : Réseaux des systèmes informatiques

Intitulée

*Contribution à la compression des images
stéréoscopiques*

Devant le jury composé de :

<i>Président :</i>	GAFOUR Abdelkader	MCA	UDL-SBA
<i>Examineurs :</i>	BELALEM Ghalem	Professeur	Université d'Oran 1
	HAMDADOU Djamila	Professeur	Université d'Oran 1
	TLEMSANI Redouane	MCA	INTTIC d'Oran
	SLAMA Zohra	MCA	UDL-SBA
<i>Directeur de thèse :</i>	FARAOUN Kamel Mohamed	Professeur	UDL-SBA
<i>Membre invité :</i>	FEZZA Sid Ahmed	MCA	INTTIC d'Oran

Année universitaire : 2018-2019

Résumé

L'image 3D a la capacité d'améliorer l'expérience multimédia des spectateurs en fournissant une sensation de profondeur supplémentaire. La plupart du contenu 3D est capturé et préparé au format stéréoscopique. La 3D stéréoscopique (S-3D) consiste à capturer deux images à partir de deux positions légèrement différentes, chacune de ces images est fournie à chaque œil du spectateur. Ainsi, le contenu stéréoscopique nécessitera deux fois plus de données que le contenu 2D monoscopique. Compte tenu de l'énorme quantité de données qu'implique la transmission des médias S-3D, le développement de techniques de compression efficaces est d'une importance primordiale. Par conséquent, cette thèse présente deux nouvelles techniques pour la compression des images stéréoscopiques. Premièrement, sur la base de la théorie de la suppression binoculaire, une image de la paire stéréo peut être codée avec un niveau de qualité inférieur à celui de l'autre image. Toutefois, à condition que l'écart de qualité entre les deux images ne dépasse pas un certain seuil. Nous proposons une nouvelle méthode qui sélectionne automatiquement l'écart de qualité maximal tolérable entre les deux images. Cela a été réalisé en modélisant la relation entre la qualité des deux vues et leur distorsion inter-vues et, également, grâce à l'inclusion du modèle inspiré du système visuel humain. La seconde technique utilise une méthode de colorisation pour la compression d'images stéréo. L'idée principale est de compresser une vue de la paire stéréo en utilisant une méthode de codage standard, alors que pour l'autre vue seule la composante de luminance est considérée pour la compression. L'information de chrominance de cette dernière vue est presque totalement ignorée, à l'exception de quelques pixels représentatifs (PRs). Ces PRs sont définis à l'aide d'une nouvelle méthode d'extraction basée sur la squelettisation. Au niveau du décodeur, les valeurs de chrominance des pixels restants sont restaurées en utilisant une méthode de colorisation. Les résultats expérimentaux ont montré que les méthodes de codage proposées permettent des gains de débit considérables et surpassent les méthodes de codage d'images stéréoscopiques de l'état de l'art.

Mots clés : Images stéréoscopiques, codage asymétrique, suppression binoculaire, BJND, système visuel humain, colorisation, transfert de couleur, pixels représentatifs.

Abstract

3D image has the capability to enhance the multimedia experience of viewers by providing an added sensation of depth. Stereoscopic 3D (S-3D) consists of capturing two images from two slightly different positions and provides each of them to each viewer's eye. Thus, stereoscopic content will require twice the amount of data compared to monoscopic one. In view of the huge amount of data involved by media transmission and consumption of such S-3D content, the development of efficient compression techniques is of paramount importance. To this end, this thesis presents two novel techniques to compress stereoscopic images. First, based on the binocular suppression theory, in order to reduce the required bitrate for stereoscopic images delivery, one image of the stereo pair can be encoded at high quality level, while the other image is encoded at a slightly lower quality, without noticeable visual quality losses. However, with the condition that the quality gap between both images does not exceed a certain threshold. We propose a novel method that selects automatically the maximum tolerable quality gap between views of different stereo image pairs. This has been achieved by modeling the relationship between the asymmetric quality and the inter-view distortion between corresponding pixels. Then, the derived relationship is combined with Binocular Just Noticeable Difference (BJND) model to control the quality threshold between the left and right views. The second technique employs colorization method for stereo image compression. The main idea is to compress one view of the stereo pair using a standard coding method, while for the other view, only the luminance component is considered for compression. The chromatic information of this latter view are transmitted to the decoder for a few representative pixels (RPs) only. These RPs are defined using a novel proposed RP extraction method based on skeletonization. At decoder side, color values of all the remaining pixels are reconstructed by colorization methods. Experimental results showed that our coding methods can achieve considerable bit-rate saving and outperforms the widely used stereoscopic images coding approaches.

Keywords : Stereoscopic images, asymmetric coding, binocular suppression, binocular just-noticeable difference, human visual system, colorization, color transfer, representative pixels.

Table des matières

Résumé	i
Abstract	ii
Remerciements	iii
Table des figures	viii
Liste des tableaux	ix
Liste d'abréviations	x
Introduction générale	1
1 Les concepts de base de la stéréovision	6
1.1 Introduction	6
1.2 Principe de base de la stéréovision	8
1.2.1 Modèle géométrique de la caméra	8
1.3 Système d'imagerie stéréoscopique	12
1.3.1 Géométrie épipolaire	12
1.3.2 La rectification épipolaire	14
1.4 Le processus d'appariement stéréoscopique	15
1.4.1 La fonction de disparité	15
1.4.2 La profondeur	16
1.4.3 Les difficultés dans le processus de l'appariement stéréoscopique	18
1.4.4 Les méthodes d'appariement stéréoscopique	19
1.5 Conclusion	22

2	Compression d'images stéréoscopiques : état de l'art	23
2.1	Introduction	23
2.2	Les outils de la compression d'image	24
2.2.1	La transformation	25
2.2.2	La quantification	28
2.2.3	Le codage entropique	30
2.2.4	Le codage progressif	30
2.3	Le codage d'images stéréoscopiques	34
2.3.1	Schéma de codage conjoint	34
2.3.2	DE/DC	37
2.3.3	La transformation	41
2.3.4	Quantification/Codage entropique	43
2.3.5	Allocation de bits	44
2.4	Conclusion	46
3	Le codage asymétrique des images stéréoscopiques	47
3.1	Introduction	47
3.2	Les méthodes de codage asymétrique des images stéréoscopiques	49
3.2.1	La quantification asymétrique	50
3.2.2	Codage par résolution mixte	51
3.2.3	Luminance/chrominance asymétrique	54
3.3	Les limites de l'asymétrie	55
3.4	La méthode de codage asymétrique proposée	55
3.4.1	Le modèle JND	58
3.4.2	Le BJND	59
3.5	Les résultats expérimentaux	62
3.6	Conclusion	67
4	Le codage des images stéréoscopiques basé sur la colorisation	68
4.1	Introduction	68
4.2	Les méthodes de codage basées sur la colorisation	70
4.3	L'approche proposée	73
4.4	Les résultats expérimentaux	76
4.5	Conclusion	78
5	Conclusion et perspectives	81

Table des matières	vi
Bibliographie	83
Publications	97

Table des figures

1.1	Principe de la chambre noire.	8
1.2	Modèle géométrique de la caméra sténopé.	9
1.3	La géométrie d'un capteur stéréoscopique.	12
1.4	La géométrie épipolaire entre une paire d'images.	13
1.5	Illustration de la rectification épipolaire	14
1.6	Processus de mise en correspondance.	15
1.7	La relation entre la disparité et la profondeur.	16
1.8	les carte de disparité et de profondeur calculées à partir de l'image stéréo <i>room3D</i>	17
1.9	Problème d'occultation.	18
1.10	Problème d'occultation	19
1.11	Méthode de mise en correspondance stéréoscopique fondée sur la cor- rélation.	20
2.1	Le schéma générique de la compression.	25
2.2	La technique DPCM dans un contexte de codage sans perte.	26
2.3	Les différentes stratégies d'ordre de balayage des pixels utilisées dans la compression d'image	29
2.4	La relation entre les coefficients dans une représentation d'image dya- dique.	32
2.5	Schéma de codage d'image stéréo	35
2.6	Cartes de disparités estimées pour l'image stéréo <i>Art</i>	41
3.1	Un exemple de quantification asymétrique.	50

3.2	Une paire d'images stéréoscopiques avec la vue de droite filtrée par un filtre passe-bas.	52
3.3	Le schéma typique du codage RM.	52
3.4	Mise en œuvre de la méthode de codage RM	53
3.5	Le schéma de l'approche proposée pour le codage asymétrique des images stéréoscopiques.	57
3.6	Les distributions de courbe de <i>InterD</i> et de <i>Q</i> à différentes valeurs de Q_l et Q_r	58
3.7	Les huit images stéréoscopiques utilisées dans l'expérience	63
3.8	Les courbes de débit-distorsion pour les huit paires d'images stéréoscopiques utilisées.	65
4.1	Le schéma proposé pour le codage d'image stéréoscopique basé sur la colorisation	73
4.2	Exemple de squelette utilisé pour extraire les PRs.	75
4.3	Courbes de débit-distorsion pour l'image gauche des paires stéréo	78
4.4	Exemples d'images gauche reconstruites	80

Liste des tableaux

3.1	Les gains de bitrate et de PSNR	66
4.1	La comparaison des gains de bitrate et des moyennes de PSNR . . .	78

Liste d'abréviations

2D	Two-Dimensional
3D	Three-Dimensional
3DTV	Three-Dimensional TeleVision
BJND	Binocular Just-Noticeable Difference
DC	Disparity Compensation
DCP	Disparity-Compensated Prediction
DCT	Discrete Cosine Transform
DPCM	Differential Pulse Code Modulation
DWT	Discrete Wavelet Transform
EBCOT	Embedded Block Coding with Optimal Truncation
EZW	Embedded Zero-trees Wavelet
ISO	International Standardization for Organization
JND	Just-Noticeable Difference
JPEG	Joint Photographic Experts Group
LOT	Lapped Orthogonal Transforms
MC	Motion Compensation
PCRD	Post-Compression Rate-Distorsion
PR	Pixel Représentatif
PSNR	Peak Signal-to-Noise Ratio
QoE	Quality of Experience
RANSAC	Random Sample Consensus
R-D	Rate-Distortion
S-3D	Stereoscopic 3D
SAD	Sum of Absolute Difference
SIFT	Scale Invariant Feature Transform
SPIHT	Partitioning In Hierarchical Trees
SSD	Sum of Squared Differences
SVH	Système Visuel Humain

Introduction générale

Le contexte

Il est incontestable que les systèmes multimédias ont amélioré la qualité de vie des êtres humains. Ces systèmes ont évolué de manière exponentielle au cours du siècle dernier pour fournir un contenu plus riche aux utilisateurs. Plusieurs de ces révolutions technologiques ont eu un grand impact, telles que le passage du noir et blanc à la couleur et du standard à la haute définition (HD), jusqu'aux derniers systèmes de télévision 3D (3DTV).

On peut remarquer que ces évolutions ont comme but commun d'offrir un contenu plus immersif et plus proche de la réalité. Les technologies immersives capturent plus d'informations que les technologies conventionnelles, elles ont la capacité de fournir plus de détails, une meilleure perception de la profondeur, ainsi qu'une meilleure couleur, contraste et luminosité. Comme le contenu immersif contient davantage d'informations que le contenu conventionnel, des algorithmes de compression efficaces sont nécessaires pour le stockage et la transmission à l'aide des infrastructures existantes.

Parmi ces technologies immersives, on peut citer les technologies d'imagerie stéréoscopique tridimensionnelle (S-3D). Ces technologies d'imagerie S-3D offrent une sensation de profondeur améliorée en présentant au spectateur deux perspectives légèrement différentes de la même scène, telles que capturées, par exemple, par deux caméras alignées horizontalement. La vue de gauche est présentée uniquement à l'œil gauche et la vue de droite est présentée à l'œil droit uniquement. Les différentes technologies S-3D diffèrent principalement sur la manière dont cette séparation fondamentale est obtenue. Les avantages de l'imagerie S-3D ne se limitent pas à une sensation de profondeur accrue, mais incluent également un sentiment

accru d'immersion, de présence et de réalisme. Les progrès récents dans les technologies d'acquisition et d'affichage ont permis l'utilisation généralisée de la S-3D dans divers domaines d'application, par exemple la téléprésence dans la vidéo-conférence et la télémédecine pour ne citer que ceux-là.

Ainsi, apporter une dimension supplémentaire de profondeur pour la perception de la scène se fait au prix d'une augmentation de la quantité de données à stocker ou à transmettre. Ces quantités de données sont prohibitives et constituent une limitation majeure pour le développement d'applications S-3D. Nominale, le contenu S-3D nécessitera deux fois plus de données que le contenu monoscopique 2D. Par conséquent, le développement de techniques de compression efficaces est d'une importance primordiale dans le stockage, la transmission et la consommation d'un tel contenu S-3D. Toutefois, une des particularités fondamentales à propos des médias S-3D réside dans le fait qu'il existe une forte corrélation entre chaque vue. Cette corrélation devra donc être exploitée par le schéma de compression, par la prédiction d'une des deux vues à partir de l'autre.

Différentes approches ont été proposées dans la littérature pour répondre à la problématique de la compression des images stéréoscopiques. Cette thèse adresse cette problématique avec pour objectif de réduire le débit tout en respectant la perception humaine du média. Parmi ces techniques prometteuses, on peut citer celles utilisées pour coder le contenu S-3D connu sous le nom de méthodes de codage asymétrique. Ces méthodes sont basées sur la théorie de la suppression binoculaire, qui stipule que si les deux vues sont de qualité différente, le système visuel humain (SVH) fusionne les deux images de sorte que la qualité binoculaire perçue pour un spectateur serait plus proche de la qualité supérieure. Cela signifie que l'une des deux vues peut être codée avec une qualité inférieure, réduisant ainsi le débit total nécessaire pour transmettre la séquence stéréoscopique sans affecter l'expérience visuelle de l'observateur.

Une question importante dans les méthodes de codage asymétrique est liée à la détermination du meilleur niveau d'asymétrie ou de la différence de qualité visuelle maximale entre les vues, garantissant ainsi une expérience de visualisation 3D confortable.

La couleur est également importante dans de nombreuses applications de l'imagerie 3D. Elle est utilisée dans la segmentation des images, ce qui permet au spectateur de faire la distinction entre les objets de l'image et entre les objets et leur arrière-plan. Il est également important de distinguer les types d'objets ou leur statut (par exemple, dans les applications médicales). Toutefois, dans la littérature,

l'information de couleur dans la compression des images stéréoscopiques est traitée comme dans le cas des images 2D monoscopiques, sans tirer parti de la grande redondance des images stéréoscopiques, car les deux images contiennent des informations sur les mêmes objets, même si elles proviennent de perspectives légèrement différentes. Peu de travaux ont été consacrés à l'amélioration de la compression de la couleur dans les images stéréoscopiques. Par conséquent, on peut s'attendre à ce que l'application de telles méthodes aux images stéréoscopiques puisse fournir une réduction significative du débit.

Contributions de cette thèse

Tout l'enjeu des travaux que nous avons entrepris durant cette thèse réside dans l'investigation de nouvelles techniques de compression des images stéréoscopiques, offrant ainsi une réduction considérable du débit tout en limitant autant que possible les dégradations perceptibles. Les principales contributions de cette thèse peuvent être énumérées comme suit :

- **Contribution 1** : Des travaux de recherche importants ont été consacrés à la technique de codage stéréoscopique asymétrique. Basé sur la théorie de la suppression binoculaire, ces méthodes permettent de coder une vue de la paire stéréo avec une qualité inférieure à celle de l'autre, ce qui permet de diminuer la bande passante nécessaire à la diffusion de contenus visuels 3D sans dégrader la qualité subjective. Cependant, à condition que l'écart de qualité entre les deux images ne dépasse pas un certain seuil. Quelques travaux ont proposé de fixer ce seuil par des expériences subjectives. Néanmoins, l'utilisation d'un seuil fixe ne permet pas l'adaptation au contenu de l'image stéréo et présente une dépendance à la conception expérimentale. Nous avons proposé dans cette thèse une nouvelle méthode qui sélectionne automatiquement l'écart de qualité maximal tolérable entre les vues pour différentes paires d'images stéréo. Ceci a été réalisé en modélisant la relation entre la qualité asymétrique et la distorsion inter-vue entre les pixels correspondants. Ensuite, la relation dérivée est combinée au modèle BJND (Binocular Just Noticeable Difference) pour contrôler le seuil de qualité entre les vues de gauche et de droite.
- **Contribution 2** : La plupart des méthodes de codage des images stéréoscopiques se sont concentrées sur la luminance et ont ignoré les informations chromatiques. Par conséquent, nous avons proposé une méthode de codage

d'image stéréoscopique en couleur basée sur la colorisation. L'idée principale est de compresser une vue de la paire stéréo en utilisant une méthode de codage standard, tandis que pour l'autre vue, seule la composante de luminance est considérée pour la compression. Les informations chromatiques de cette dernière vue sont transmises au décodeur uniquement pour quelques pixels représentatifs qui sont définis en utilisant une nouvelle méthode d'extraction basée sur la squelettisation. Au niveau du décodeur, les valeurs de couleur de tous les pixels restants sont reconstruites par des méthodes de colorisation.

Organisation

Cette thèse est organisée en quatre chapitres, deux chapitres sont consacrés à la présentation et l'analyse de la littérature existante liée aux travaux présentés dans cette thèse, deux chapitres correspondant à chacune des contributions et un autre consacré aux conclusions et travaux futurs.

Dans le **chapitre 1**, nous introduisons d'abord les concepts de base des modèles de caméra, puis nous établissons les relations géométriques entre les deux vues d'une paire stéréoscopique. Nous passons également en revue les techniques les plus importantes visant à estimer le champ de disparité.

Au **chapitre 2**, en premier lieu, un aperçu des schémas de codage d'images stéréoscopiques est donné. Ensuite nous présentons les concepts les plus importants derrière les techniques de compression d'images. Ensuite un état de l'art sur les travaux antérieurs pour chaque étape du processus de codage d'images stéréoscopiques est donné.

Le **chapitre 3** est dédié à notre première contribution qui est consacrée à la problématique du codage asymétrique des images stéréoscopiques. L'accent est mis dans ce chapitre sur l'identification du seuil d'asymétrie juste perceptible entre les vues de la paire stéréo pour lequel l'effet 3D stéréoscopique n'est pas affecté.

Chapitre 4 présente notre deuxième contribution portant sur le codage par colorisation des images stéréoscopiques. Cette approche vise à mieux compresser l'information chromatique, et ceci en considérant seulement la composante de luminance pour la compression d'une seule vue de la paire stéréo. Les informations chro-

matiques de cette vue sont transférées de l'image de référence (c'est-à-dire l'image couleur) vers l'image cible.

Le dernier chapitre, le **chapitre 5**, résume les contributions de cette thèse et donne des orientations futures.

Le codage asymétrique des images stéréoscopiques

Sommaire

3.1	Introduction	47
3.2	Les méthodes de codage asymétrique des images stéréoscopiques	49
3.2.1	La quantification asymétrique	50
3.2.2	Codage par résolution mixte	51
3.2.3	Luminance/chrominance asymétrique	54
3.3	Les limites de l'asymétrie	55
3.4	La méthode de codage asymétrique proposée	55
3.4.1	Le modèle JND	58
3.4.2	Le BJND	59
3.5	Les résultats expérimentaux	62
3.6	Conclusion	67

3.1 Introduction

Ces dernières années, le contenu multimédia produit est très riche et les systèmes multimédia ont évolué de manière exponentielle. Ces systèmes visent à améliorer la qualité d'expérience (QoE) des utilisateurs en leur offrant un sentiment réaliste

d'immersion. L'exemple le plus évident est celui des applications de médias visuels en 3D qui donnent l'illusion d'une perception de profondeur de la scène. Il existe différentes techniques pour représenter le contenu 3D, mais la plupart du contenu 3D est capturé et préparé au format stéréoscopique [78]. La 3D stéréoscopique (S-3D) consiste en deux vues de la même scène capturées par deux caméras séparées par une baseline spécifique. Ainsi, le contenu stéréoscopique nécessitera deux fois plus de données que le contenu monoscopique. Cette augmentation est une question importante ayant attiré un nombre important d'efforts de recherche. Le développement de techniques de compression efficaces est d'une importance primordiale dans le stockage, la transmission multimédia et la consommation d'un tel contenu S-3D.

Le moyen le plus évident et facile pour coder les images stéréoscopiques est d'encoder indépendamment chaque vue de la paire stéréo, en utilisant n'importe quelle norme de compression d'image fixe, comme JPEG [79] ou JPEG 2000 [80]. Étant donné que l'observateur humain étant le récepteur ultime (end user), il serait avantageux d'intégrer les propriétés du système visuel humain (SVH) pour améliorer les performances de codage. Au cours de la dernière décennie, nous pouvons constater une augmentation considérable de l'utilisation des approches perceptuelles pour améliorer les performances des techniques de traitement multimédia [52]. Sur la base des caractéristiques du SVH, différentes méthodes ont été développées dans de nombreux domaines complexes du traitement de l'image tels que le codage d'image [82], l'amélioration et le tatouage, l'évaluation de la qualité d'image, etc. De cette façon, certaines approches perceptuelles exploitant les propriétés perceptuelles du SVH ont été proposées pour le codage d'images stéréoscopiques.

Parmi ces techniques prometteuses, on peut citer celles utilisées pour coder le contenu S-3D, connu sous le nom de méthodes de codage asymétrique. Les techniques de codage asymétrique sont basées sur la théorie de suppression binoculaire [83–86], qui spécifie que si les deux yeux sont fournis avec deux vues de qualité différente, l'image 3D perceptive fusionnée est dominée par la vue de qualité supérieure [81–83], [87–90]. Par conséquent, selon cette théorie, afin de réduire le débit binaire requis pour la paire d'images stéréoscopiques, une vue de la paire peut être codée avec une qualité suffisamment élevée, tant dis que l'autre vue est codée avec une qualité légèrement inférieure (avec un bit-budget inférieur), et ceci sans que l'utilisateur ne s'aperçoit de cela.

Cette approche a concentré des efforts de recherche considérables et elle a été exploitée avec différentes manières. Selon la manière dont la qualité (information) d'une des vues est réduite, les différentes méthodes appartenant à ce concept

peuvent être classées en quatre catégories : (a) réduction de la résolution spatiale (filtrage spatial), (b) quantification asymétrique (QP inégale), (c) réduction de la résolution et (d) luminance/chrominance asymétrique.

Un obstacle important dans les méthodes de codage asymétrique des images stéréoscopiques est lié au niveau d'asymétrie entre les vues, c'est-à-dire le seuil d'asymétrie tout juste visible, de sorte que l'expérience de visionnement en 3D n'est pas altérée. Dans le but de résoudre ce problème, plusieurs travaux ont tenté d'identifier ce seuil d'asymétrie avec des expériences subjectives. Même si ces études ont réussi à définir ou à identifier des seuils, ceux-ci dépendent fortement de la conception de l'expérience et ne permettent aucune adaptation aux variations de la qualité et du contenu de l'image. Par conséquent, dans ce chapitre, nous proposons une nouvelle méthode de codage asymétrique [143] ajustant de manière adaptative le niveau d'asymétrie pour chaque image afin d'assurer une bonne perception binoculaire. Ceci est réalisé en exploitant plusieurs modèles inspirés de SVH.

Le reste du chapitre est organisé comme suit : La section 3.2 décrit les différents types d'asymétrie qui ont été proposés dans la littérature pour le codage des images stéréoscopiques. Un survol des différentes tentatives les plus importantes pour la détermination des limites perceptives de l'asymétrie dans le contexte du codage des images stéréoscopiques est présenté à la section 3.3. La section 3.4 décrit l'approche proposée et les différents outils exploités. Les résultats expérimentaux sont présentés à la section 3.5. La section 3.6 résume et conclut ce chapitre.

3.2 Les méthodes de codage asymétrique des images stéréoscopiques

Basé sur la théorie de la suppression binoculaire, le SVH fusionne les images rétinienne provenant des deux yeux ayant différents niveaux de qualité de telle sorte que la qualité visuelle perçue est proche de la vue avec une qualité supérieure. Cette propriété est effectivement exploitée pour réduire le débit binaire requis pour la transmission des images stéréoscopiques, où une vue de la paire stéréo est codée fortement par rapport à l'autre vue sans induire d'artefacts visibles dans la perception stéréoscopique.

Sur la base de cette découverte, de nombreuses méthodes de codage asymétriques sont apparues dans la littérature. En fonction de la manière dont l'asymétrie de qualité entre les deux vues peut être réalisés, les différentes méthodes de codage

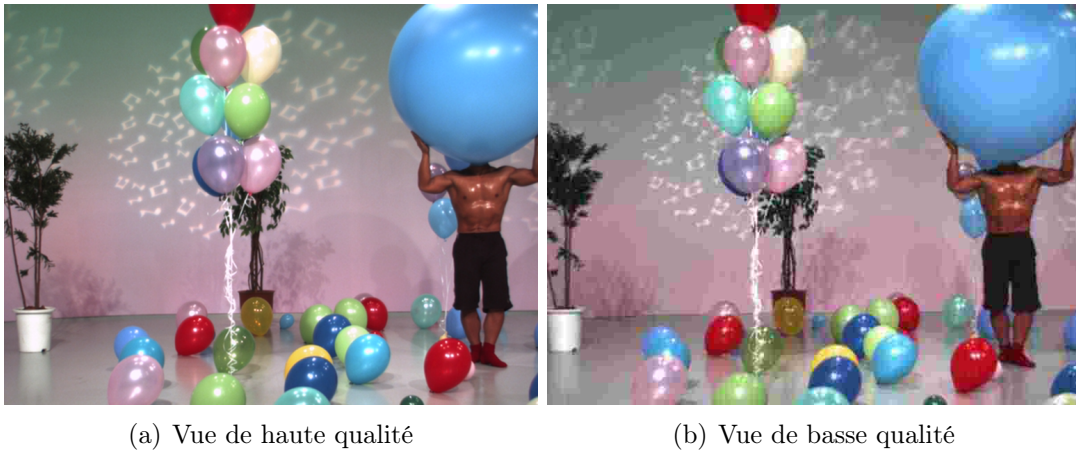


FIGURE 3.1 – Un exemple de quantification asymétrique.

asymétrique des images stéréoscopiques peuvent être classées en trois catégories principales détaillées ci-après.

3.2.1 La quantification asymétrique

Une des méthodes pour réduire la bande passante requise pour la transmission d'images stéréoscopiques consiste à coder les deux images de la paire stéréo avec des PSNR ou des débits binaires différents, cette méthode est généralement connue sous le nom de réduction de PSNR (qualité) ou de quantification asymétrique. La mise en œuvre de l'asymétrie par réduction du PSNR est simple et ne nécessite aucune modification du codeur/décodeur. Étant donné que la qualité d'encodage d'une vue dépend du paramètre de quantification, l'utilisation de différents paramètres de quantification pour les vues de gauche et de droite conduit à une asymétrie de la qualité. Par conséquent, une vue est codée avec une plus grande valeur de quantification que l'autre. Ce qui résulte en une paire stéréo avec une des vues présentant une plus grande quantité d'artefacts de codage visibles, tels que le flou ou l'effet de bloc. Un exemple de l'effet de la quantification asymétrique est présenté sur la paire stéréo de la figure 3.1.

L'effet de la méthode de quantification asymétrique a été étudié dans [98]. Seun-tiens et al. ont appliqué un taux de compression différent sur les vues gauche et droite d'une paire stéréo en utilisant JPEG [117]. Les résultats ont montré qu'un codage JPEG asymétrique n'avait aucun effet sur la profondeur perçue. Par ailleurs les résultats sur le codage asymétrique et symétrique ont montré que la relation

entre la qualité perçue et le débit binaire moyen n'est pas simple. Dans certains cas, l'évaluation de la qualité d'image d'une paire codée de manière symétrique peuvent être plus élevées que pour une paire codée de manière asymétrique, même si le débit binaire moyen pour la paire symétrique est plus faible que pour la paire asymétrique. Cette conclusion a été aussi rapportée par les travaux décrits dans [98].

Saygili *et al.* [92] ont proposé d'encoder une vue de la paire stéréo en utilisant le codeur H.264 à une qualité suffisamment élevée (PSNR = 40 dB), tandis que l'autre vue est codée avec une qualité inférieur, mais maintenue au-dessus d'un certain seuil. Ils ont remarqué que la dégradation en visualisation 3D est imperceptible. Les résultats indiquent qu'à des débits élevés, le codage asymétrique utilisant la méthode de réduction du PSNR permet d'obtenir une bonne qualité perceptuelle. Cependant, à de faibles débits, le codage symétrique produit une meilleure qualité, ce qui est conforme à la conclusion de Seuntiens *et al.* dans [46].

Enfin, pour fournir une qualité 3D acceptable en exploitant ce type d'asymétrie, un compromis entre la dégradation de la qualité et la réduction du débit doit être envisagé. Ce point a fait l'objet de notre contribution décrite à la section 3.4.

3.2.2 Codage par résolution mixte

Selon le phénomène perceptif de la suppression binoculaire, le SVH peut tolérer l'absence d'informations à haute fréquence dans une seule vue. Il est possible de réduire la résolution spatiale d'une vue de la paire stéréoscopique tout en gardant l'autre vue à sa haute résolution d'origine, et sans affecter la perception 3D globale. Cette approche est généralement appelée méthode de résolution spatiale mixte (RM). Par conséquent, dans le concept de la RM, les images gauche et droite sont présentées au spectateur avec différents niveaux de netteté, c'est-à-dire qu'une image de la paire stéréo est floue, tandis que la seconde est conservée telle quelle (voir figure 3.2). Des tests subjectifs ont démontré que la fusion 3D perçue de ces deux images est relativement proche de l'image la plus nette [99]. Cela peut s'expliquer par le fait que l'image nette masque l'effet de flou de la qualité inférieure.

Perkins a été le premier à introduire la notion de codage RM [119]. Afin de réduire le débit pour la transmission S-3D, avant l'étape de codage, Perkins a proposé d'appliquer un filtrage passe-bas à une seule vue de la paire stéréo, tandis que la deuxième vue reste telle qu'elle. Du fait qu'il y a moins de pixels à coder par rapport au cas de la résolution complète, cette opération réduit considérablement

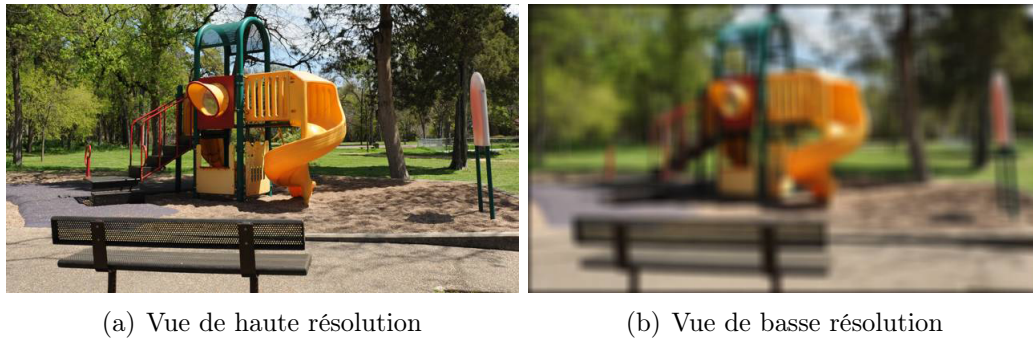


FIGURE 3.2 – Une paire d’images stéréoscopiques avec la vue de droite filtrée par un filtre passe-bas.

le débit binaire. Au niveau du décodage, la vue avec la résolution la plus faible est filtrée par un filtre passe-haut afin de correspondre à la résolution de la seconde vue, comme illustré à la figure 3.3. Plusieurs travaux ont adopté ce concept depuis lors [104–113]. En ce qui concerne la littérature, le codage RM est de loin le plus exploré et le plus couramment utilisé parmi les approches de codage asymétrique.

Dans [91], basé sur des tests subjectifs, les auteurs ont étudié l’impact du codage RM avec des rapports différents de sous-échantillonnage ($1/2$ et $1/4$). Il a été constaté que le filtrage spatial produisait des résultats acceptables et n’avait aucun effet sur la sensation globale de la profondeur. La comparaison des performances de la quantification asymétrique et des méthodes à base de RM ont été conduites dans [98, 99, 105, 106, 111, 112].

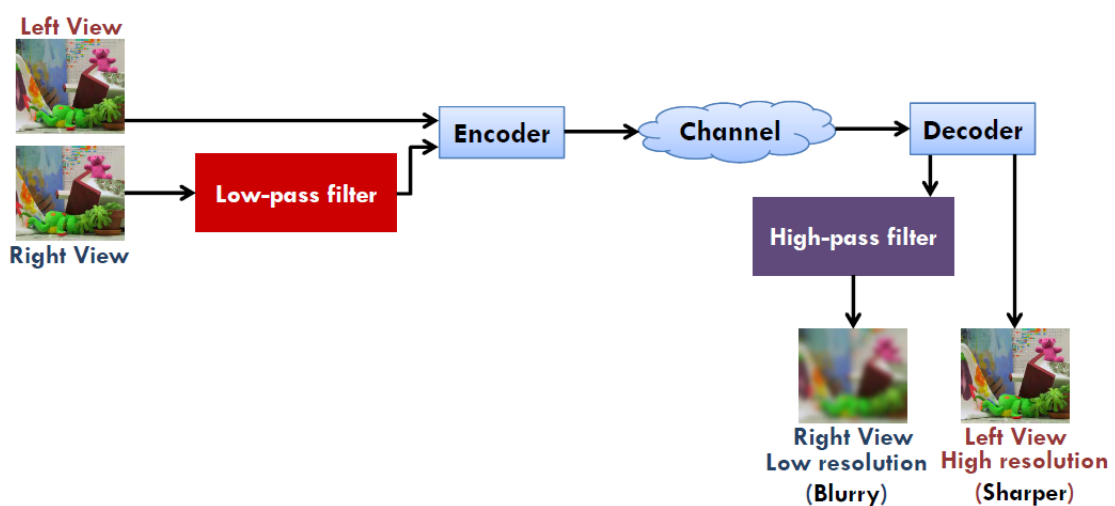
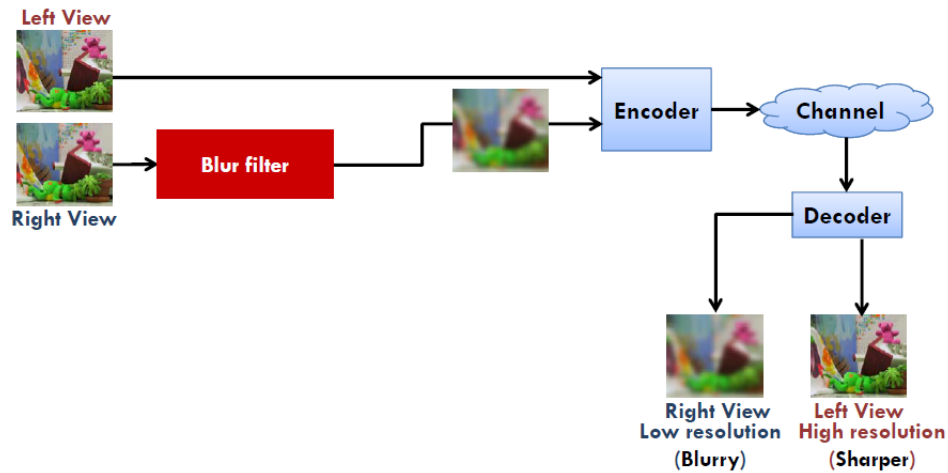
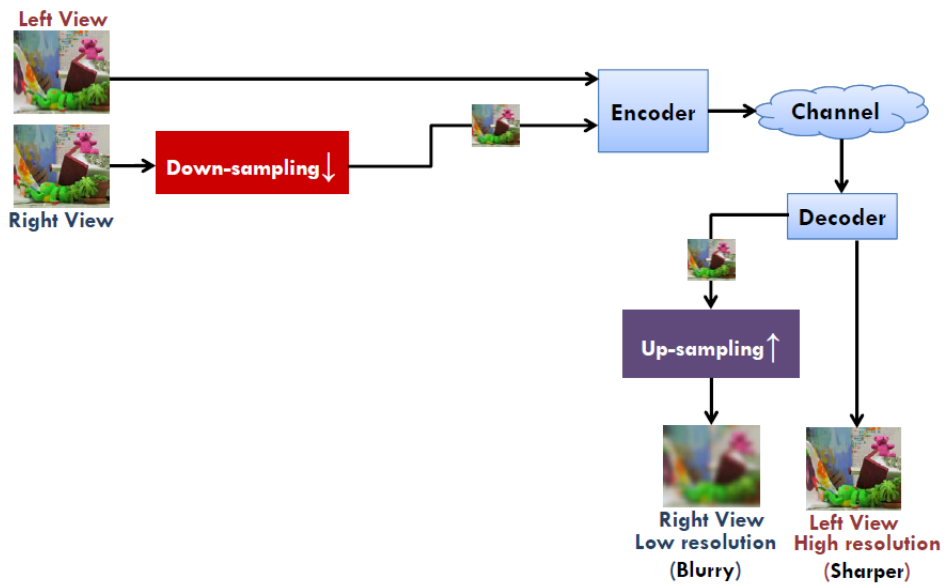


FIGURE 3.3 – Le schéma typique du codage RM.



(a)



(b)

FIGURE 3.4 – Mise en œuvre de la méthode de codage RM par (a) un filtre de flou ou (b) des approches sur-/sous- échantillonnage.

Généralement, il existe deux façons d'obtenir le codage RM. La première est un sous-échantillonnage d'une image à une taille plus petite avant le codage, et la deuxième applique un filtre de flou (par exemple, un filtre à disque, un filtre gaussien, etc.), comme illustré sur la figure 3.4. Dans le domaine spatial, le sous-échantillonnage d'une image réduit sa taille, car seuls des pixels spécifiques sont conservés en fonction du rapport utilisé. Cependant, le filtre de flou supprime les

composants haute fréquence, tandis que la taille de l'image reste inchangée. Cela permet de coder l'image avec moins de bits, car seuls quelques détails (composants haute fréquence) doivent être codés. Du point de vue de la qualité visuelle, les deux procédures ont le même effet visuel de flou. Cependant, le sous-échantillonnage, puis le sur-échantillonnage pour la visualisation 3D, introduit plus d'artefacts visuels que l'utilisation de filtres de flou [110,111]. En général, on peut conclure que le codage RM est une approche prometteuse pour réduire le débit binaire tout en fournissant une qualité visuelle 3D acceptable.

3.2.3 Luminance/chrominance asymétrique

Inspiré par des travaux antérieurs sur la perception visuelle, rapportant que le SVH est beaucoup plus sensible à la luminance qu'à la chrominance et que, l'acuité visuelle est moindre pour la chrominance par rapport à la luminance. Certains travaux ont proposé de représenter les informations de chrominance d'une vue de la paire stéréo par moins d'échantillons que leurs correspondants dans l'autre vue [114]. Il a été montré dans [114] que l'application d'un sous-échantillonnage aux composantes de chrominance d'une seule vue de la paire stéréo n'affecte pas la qualité perçue par rapport à l'approche symétrique, à condition que le rapport de sous-échantillonnage soit correctement sélectionné.

D'autres travaux ont proposé d'ignorer complètement les informations de chrominance d'une vue de la paire stéréo avant le codage [115,116]. Après le décodage, l'image sans information de chrominance est colorisée en appliquant des techniques de transfert de couleur basées sur l'autre vue et l'information de disparité. Les différences entre les images reconstruites et originales sont insignifiantes, et le SVH peut ignorer ces différences en raison de la suppression binoculaire. Ainsi, les exigences de bande passante peuvent être assouplies sans affecter l'effet stéréoscopique final.

Enfin, les méthodes de codage stéréoscopique asymétrique précédemment décrites peuvent être combinées, avec la condition de garder une vue complètement inchangée. Par exemple, une combinaison de réduction de la résolution spatiale, de réduction de la chrominance et/ou d'augmentation de la QP peut conduire à une meilleure efficacité.

3.3 Les limites de l'asymétrie

Alors que des expériences subjectives ont montré que le codage asymétrique des images stéréoscopiques fournit une expérience visuelle acceptable, la quantité de cette asymétrie reste un problème ouvert. La mesure des limites perceptuelles du codage asymétrique ou la quantification du seuil d'asymétrie tout juste perceptible entre les vues de la paire stéréo, pour laquelle l'effet S-3D n'est pas affecté, est un problème important.

À notre connaissance, seuls quelques travaux ont été consacrés à ce sujet [92, 104, 105, 115]. Par exemple, les auteurs de [92] ont trouvé que le niveau d'asymétrie juste visible dépendait de la technologie d'affichage et ont exprimé cet écart entre les deux vues en terme de PSNR, soit environ 9 dB pour un écran de parallaxe et 7 dB pour une projection pleine résolution. Dans le même axe, Shao *et al.* [115] ont montré que 2 dB de différence entre les vues de gauche et de droite, est le niveau de tolérance maximal pour lequel le spectateur ne remarque pas les artefacts.

Concernant la limite du codage asymétrique par réduction de la résolution spatiale, Aflaki *et al* [104] ont mené plusieurs essais avec différents rapports de sous-échantillonnage ($1/2$, $3/8$ et $1/4$) le long des axes horizontaux et verticaux, et ont conclu que le sous-échantillonnage entre $1/2$ et $3/8$ fournit une expérience de visionnage 3D satisfaisante.

Les dernières études décrites ci-dessus, utilisent la limite d'asymétrie définie pour effectuer un codage asymétrique efficace de telle sorte que le spectateur ne puisse percevoir aucune dégradation. Cependant, tous les travaux antérieurs sont basés sur des expériences subjectives et la détermination du niveau d'asymétrie repose fortement sur eux. Malheureusement, il est difficile de généraliser le seuil dérivé à toutes autres paires stéréo, car ceci dépend du contenu de l'image, du type d'affichage, de la taille de l'écran et ainsi de suite. Une démonstration de cette variabilité peut être observée entre [92] et [115], où les deux études se sont concentrées sur la quantification asymétrique suivant la même procédure, les résultats ont donné des valeurs de seuils très différentes.

3.4 La méthode de codage asymétrique proposée

Dans ce chapitre, nous proposons une nouvelle méthode de codage asymétrique qui détermine automatiquement l'écart de qualité le plus approprié entre les images de gauche et de droite d'une paire stéréoscopique [143], ce qui offre une meilleure

expérience de visualisation 3D. Pour réaliser ce contrôle de l'asymétrie, la relation entre la qualité asymétrique et la distorsion inter-vues entre les pixels correspondants, est exprimée sous forme d'une fonction reposant sur le modèle perceptuel BJND (Binocular Just-Noticeable Difference) [95].

Comme il a été indiqué précédemment, lorsque les images de gauche et de droite sont présentées au spectateur avec différents niveaux de qualité, grâce au phénomène de suppression binoculaire du SVH, la qualité de la perception en 3D est dominée par la vue de haute qualité. Par conséquent, on se basant sur ce concept, nous pouvons compresser une vue de la paire stéréo plus fortement que l'autre sans introduire d'artefacts visibles pendant la visualisation 3D, à condition que la différence de la qualité entre les deux vues ne dépasse pas un seuil donné.

Le schéma général de la méthode proposée est présenté sur la figure 3.5. Soit une paire stéréo composée d'une image gauche I_l et une image droite I_r . Nous considérons I_l comme image de référence et I_r comme image cible. Afin de réduire le bitrate, I_r est l'image que nous allons compresser fortement et qui aura donc la qualité la plus faible.

Pour pouvoir ajuster la qualité asymétrique de la paire stéréo, c'est-à-dire ajuster la qualité de I_r (Q_r) par rapport à celle de I_l (Q_l), il est important de modéliser la relation entre la qualité des deux vues (Q) et leur distorsion inter-vue ($InterD$). On suppose que la relation $InterD$ - Q peut être décrite par une fonction comme suit :

$$InterD = f(Q) \quad (3.1)$$

où Q représente le rapport entre les qualités des deux vues ($\frac{Q_r}{Q_l}$), cette dernière prend ces valeurs dans $[0, 1]$. Alors que $InterD$ représente la distorsion entre les pixels correspondants des deux vues. Cette distorsion est calculée par la mesure d'erreur quadratique moyenne (MSE) comme suit :

$$InterD = \frac{1}{n} \sum_{i=1}^n (I_l(x, y) - I_r(x, y + d_i))^2 \quad (3.2)$$

où n est le nombre de pixels correspondants entre I_l et I_r . $I_l(x, y)$ indique la valeur de luminance du pixel i à la position (x, y) . d_i est la valeur de la disparité du pixel i .

Ensuite, afin de définir la fonction $f(\cdot)$, la relation $InterD$ - Q est évaluée expérimentalement à différentes valeurs en utilisant un ensemble d'images stéréoscopiques différentes [127]. Pour chaque paire stéréo, nous avons utilisé dix différentes valeurs

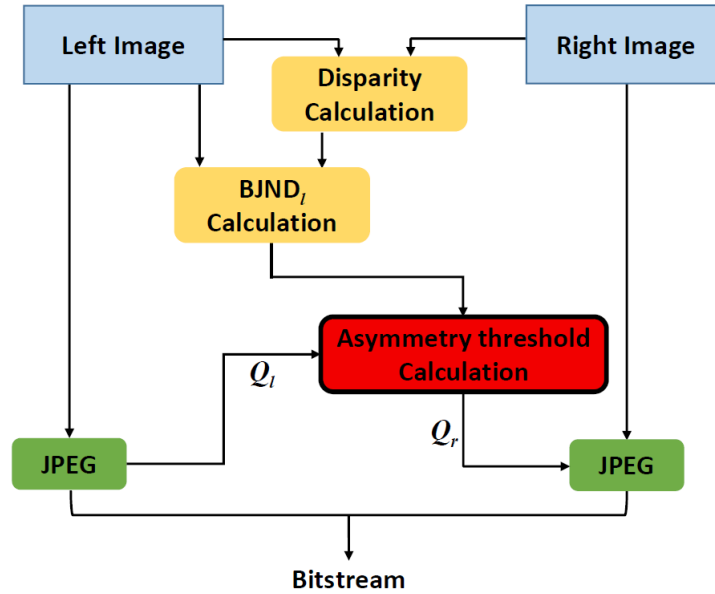


FIGURE 3.5 – Le schéma de l’approche proposée pour le codage asymétrique des images stéréoscopiques.

de Q_l allant de 10 à 100, avec un pas de 10, et nous avons fait varier les valeurs de Q_r sur la base de Q_l comme suit :

$$Q_r = Q_l - i, \quad \text{with } i \in [0, Q_l]. \quad (3.3)$$

La figure 3.6 montre les distributions de courbes de $InterD$ et Q pour différentes valeurs Q_l et Q_r .

Sur la base des résultats obtenus, nous pouvons observer que la forme générale de la distribution suit une courbe exponentielle. Par conséquent, la fonction décrivant la relation $InterD-Q$ définie par l’équation (3.1) peut être reformulée comme suit :

$$InterD = \alpha_1 \exp\left(-\alpha_2 \cdot \frac{Q_r}{Q_l}\right) \quad (3.4)$$

où α_1 et α_2 sont des paramètres du modèle et qui sont déterminés expérimentalement.

Comme mentionné précédemment, l’objectif est la réduction de la qualité d’une seule vue sans introduire d’artefacts visibles dans la visualisation 3D. En d’autres termes, la diminution de Q_r devrait conduire à un changement en dessous du seuil de visibilité binoculaire (BJND) [95]. Dans ce qui suit, nous donnons un aperçu sur le modèle BJND.

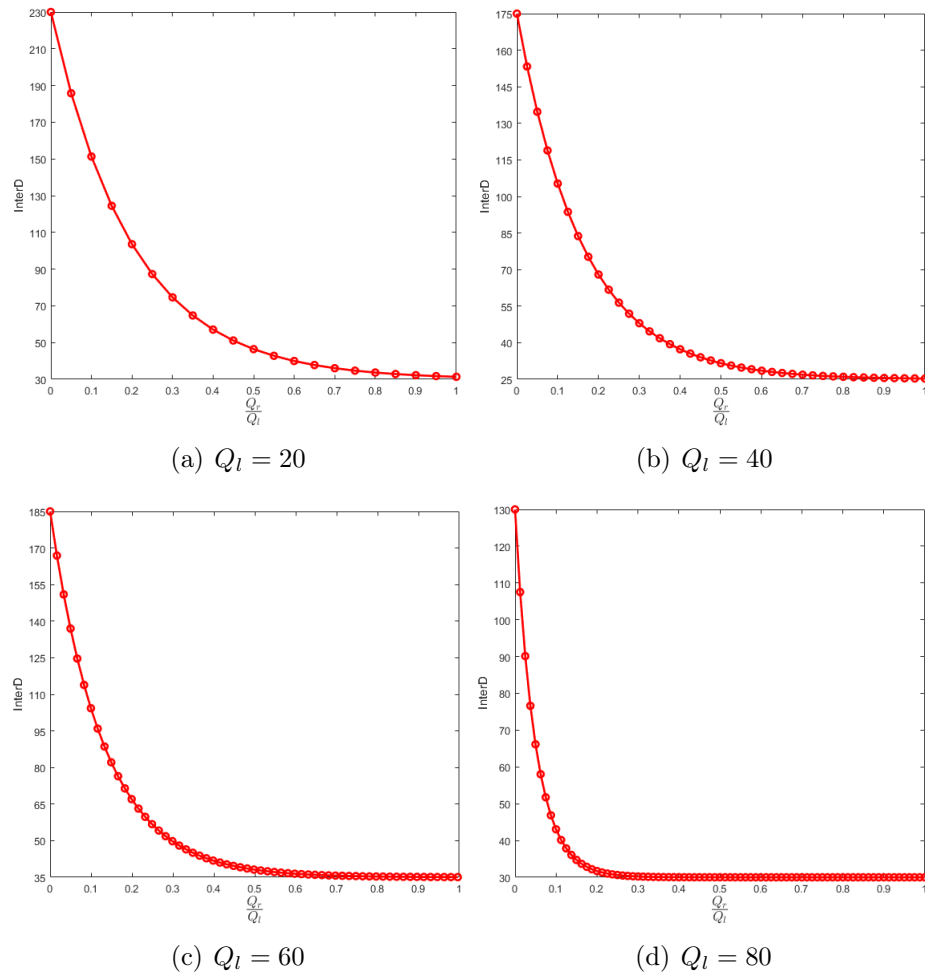


FIGURE 3.6 – Les distributions de courbe de $InterD$ et de Q à différentes valeurs de Q_l et Q_r , où Q_r et $InterD$ sont calculés respectivement par les équations (3.3) et (3.2).

3.4.1 Le modèle JND

Dans la littérature, plusieurs modèles JND (Just Noticeable Difference) ont été proposés. Typiquement, ils peuvent être déterminés dans le domaine spatial ou dans le domaine fréquentiel, tels que la transformée en cosinus discrète (DCT) et la transformée en ondelettes discrètes (DWT), ou la combinaison des deux [103].

Les modèles JND dans le domaine spatial [100, 102], nommés JND à base de pixels, se concentrent principalement sur l'adaptation de la luminance de fond et le masquage de contraste spatial. Par exemple, le modèle proposé par Chou *et al.* prend en compte l'adaptation de la luminance et le masquage de texture [102]. Ce

dernier est déterminé en fonction des activités spatiales locales de la région (par exemple, caractérisé par le calcul de gradients autour du pixel). Enfin, le JND est considéré comme l'effet dominant entre le masquage de texture et le masquage de luminance. Ce modèle a été amélioré par Yang *et al.* dans [97].

Cependant, les modèles JND basés sur les pixels ne prennent pas en compte les sensibilités de la vision humaine telles que le CSF (contrast sensitivity function). Par conséquent, il ne peut pas décrire les propriétés du SVH avec précision. Pour surmonter ce problème, certains auteurs ont proposé de concevoir des modèles JND dans le domaine fréquentiel. Les modèles JND générés dans le domaine fréquentiel, à savoir le JND basé sur la sous-bande, intègrent généralement tous les principaux facteurs affectant, tels que le CSF, l'adaptation de luminance et le masquage de contraste. L'un des modèles basé sur la DCT le plus ancien et le plus connu est celui développé par Watson *et al.* [96]. Puisque la plupart des normes de compression d'image et de vidéo sont basées sur le domaine DCT, ce modèle a été adopté dans de nombreux travaux. Dans ce modèle, le seuil JND pour chaque sous-bande DCT est calculé en incorporant les mécanismes les plus importants du SVH cités précédemment.

Plus récemment, Wei *et al.* [94] incorporent de nouvelles formules d'adaptation de luminance, de masquage de contraste et de correction gamma pour estimer le seuil JND dans le domaine de la DCT. Zhang *et al.* [103] proposent d'estimer le JND en additionnant les effets dans la DCT et dans le domaine spatial.

Concernant les modèles JND dans le domaine des ondelettes, Liu *et al.* ont proposé un modèle JND pour la transformée en ondelettes discrète [93]. Les fonctions d'adaptation de luminance et de masquage de contraste utilisées dans ce modèle sont les mêmes que dans le modèle de Watson (avec les mêmes paramètres). Pour la modélisation CSF, ils utilisent les seuils de détection d'ondelettes mesurés dans les expériences de Watson *et al.* [96].

3.4.2 Le BJND

Afin de construire un modèle de JND binoculaire (BJND pour Binocular Just-Noticeable Difference), Zhao *et al.* ont conduit des expériences psychophysiques [95]. Sur la base de ces expériences, ils ont proposé un modèle BJND pour mesurer la visibilité conjointe d'une paire de distorsions dans deux vues. Le modèle BJND détermine les distorsions minimales dans une vue qui provoquent des différences binoculaires visibles, en tenant en compte des informations d'arrière-plan et des

distorsions dans la zone correspondante de l'autre vue. Dans leur modèle proposé, ils ont considéré deux caractéristiques du SVH, à savoir les effets de masquage de luminance et de contraste, et ils les ont modélisés pour le cas de la vision binoculaire comme décrit ci-après. Soient les vues de gauche et de droite, ainsi que la carte de disparité de l'image gauche. Le BJND de l'image gauche ($BJND_l$) est défini par :

$$\begin{aligned} BJND_l(i, j) &= BJND_l(bg_r(i - d_l, j), eh_r(i - d_l, j), n_r(i - d_l, j)) \\ &= A_C(bg_r(i - d_l, j), eh_r(i - d_l, j)) \\ &\quad \times \left(1 - \left(\frac{n_r(i - d_l, j)}{A_C(bg_r(i - d_l, j), eh_r(i - d_l, j))} \right)^\lambda \right)^{\frac{1}{\lambda}} \end{aligned} \quad (3.5)$$

où i et j sont les coordonnées des pixels, d_l est la valeur de la disparité horizontale du pixel (i, j) . Le paramètre λ contrôle l'influence du bruit dans l'image de droite, et il a été suggéré dans [95] que $\lambda = 1.25$. Il convient de noter que $BJND_l$ dépend de la luminance de fond bg_r , de la hauteur du contour eh_r et de l'amplitude de bruit n_r du pixel correspondant dans l'image de droite. S'il n'y a aucun bruit dans l'image de droite, c'est-à-dire $n_r(i - d, j) = 0$, le $BJND_l$ est réduit à A_C , ce qui est défini par :

$$A_C(bg, eh) = A_{limit}(bg) + K(bg) \cdot eh \quad (3.6)$$

Grâce à des expériences psychophysiques, les auteurs ont défini $A_{limit}(bg)$ et $K(bg)$, respectivement par :

$$A_{limit}(bg) = \begin{cases} 0.0027 \cdot (bg^2 - 96 \cdot bg) + 8, & \text{if } 0 \leq bg < 48 \\ 0.0001 \cdot (bg^2 - 32 \cdot bg) + 1.7, & \text{if } 48 \leq bg \leq 255 \end{cases} \quad (3.7)$$

$$K(bg) = -10^{-6} \cdot (0.7 \cdot bg^2 + 32 \cdot bg) + 0.07 \quad (3.8)$$

où bg est la moyenne des valeurs de luminance d'un bloc de 5×5 centré sur la position du pixel correspondant, et la hauteur du contour eh est calculée par les opérateurs 5×5 du filtre Sobel comme suit :

$$eh(i, j) = \sqrt{E_H^2(i, j) + E_V^2(i, j)}, \quad (3.9)$$

$$E_k(i, j) = \frac{1}{24} \sum_{h=1}^5 \sum_{v=1}^5 I(i-3+h, j-3+v) \cdot G_k(h, v), \quad (3.10)$$

avec $k = H, V$.

où la représentation détaillée de l'opérateur Sobel $G_k(h, v)$ est comme suit :

$$G_H = \begin{pmatrix} -1 & -2 & 0 & 2 & 1 \\ -2 & -3 & 0 & 3 & 2 \\ -3 & -5 & 0 & 5 & 3 \\ -2 & -3 & 0 & 3 & 2 \\ -1 & -2 & 0 & 2 & 1 \end{pmatrix}, \quad G_V = \begin{pmatrix} 1 & 2 & 3 & 2 & 1 \\ 2 & 3 & 5 & 3 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ -2 & -3 & -5 & -3 & -2 \\ -1 & -2 & -3 & -2 & -1 \end{pmatrix} \quad (3.11)$$

Il est important de noter qu'il s'agit du premier modèle qui mesure le seuil de distorsion perceptible des images stéréoscopiques pour une vision binoculaire, et nous l'avons exploité dans la méthode proposée pour assurer une QoE 3D fiable.

Ainsi, le contrôle de la distorsion est réalisé dans la méthode proposée en limitant la distorsion inter-vue (c'est-à-dire *InterD*) en utilisant le modèle BJND comme suit :

$$InterD < BJND \quad (3.12)$$

Par conséquent, si nous substituons *InterD* dans l'équation (3.12) par sa définition dans (3.4), nous pouvons ainsi dériver automatiquement la valeur appropriée pour Q_r :

$$Q_r > -\frac{Q_l}{\alpha_2} \cdot \ln \left(\frac{BJND}{\alpha_1} \right) \quad (3.13)$$

De cette manière, pour éviter toute distorsion visible lors de l'exploitation de la méthode de codage asymétrique, la valeur sélectionnée de Q_r ne doit pas être inférieure au seuil défini en (3.13). En d'autres termes, en fonction de la valeur Q_l en plus du *BJND* qui est calculée par rapport au contenu des deux vues, nous pouvons calculer automatiquement la valeur Q_r optimale qui permet une réduction de débit significative tout en offrant une qualité d'expérience optimale.

3.5 Les résultats expérimentaux

Dans cette section, les résultats de simulation sont fournis pour évaluer la performance de la méthode de codage asymétrique proposée pour les images stéréoscopiques. Les évaluations ont été exécutées sur la base de données stéréoscopiques appelée *Middlebury* [127]. Nous avons utilisé huit paires d'images stéréo avec différentes résolutions allant de 450×375 à 1390×1110 , nommées *Art*, *Livres*, *Cloth3*, *Cônes*, *Midd1*, *Midd2*, *Moebius* et *Teddy*. La figure 3.7 montre l'image de gauche de chaque paire stéréo utilisée dans l'évaluation des performances.

La méthode proposée dépend des paramètres α_1 et α_2 (voir l'équation 3.13). Ces derniers ont été déterminés par une expérimentation approfondie et, en conséquence, les résultats optimaux ont été obtenus en fixant α_1 à 25 et α_2 à 0.35, étant donné que les valeurs du BJND ne dépassent pas 20.

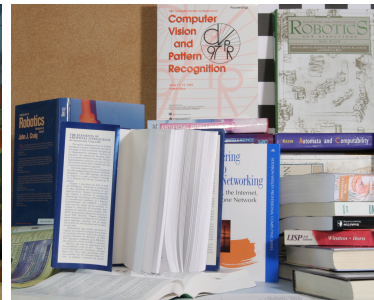
La méthode proposée a été comparée à la méthode de codage symétrique des images stéréoscopiques, qui est considérée comme la méthode de codage de référence (notée *SC*). Pour cette méthode, les deux images de la paire stéréo ont été codées de la même manière, sans privilégier une image au détriment de l'autre. De plus, trois méthodes de codage asymétrique qui sont largement utilisées ont également été considérées dans la comparaison.

La première méthode réalise l'asymétrie en appliquant sur une image un sous-échantillonnage avec un rapport de 1/2 le long des coordonnées des deux axes (notée *AC_Down*). La seconde consiste à flouter une vue de la paire stéréo en utilisant un filtre de flou gaussien (notée *AC_Blur*) et, enfin, dans la troisième méthode, les deux vues sont codées avec une qualité inégale, c'est-à-dire avec un écart de qualité fixe de 10 (noté *AC_FQualité*).

Pour toutes ces méthodes de codage, nous avons utilisé la norme JPEG [79] comme standard de codage, avec dix valeurs de qualité différentes allant de 10 à 100. Comme mentionné précédemment, l'image de gauche est codée comme image de référence et l'image de droite comme image cible. Les résultats sont présentés sous la forme d'une comparaison quantitative en termes de débit moyen et en termes de qualité PSNR correspondante. Cette dernière mesure est calculée comme une moyenne de l'erreur quadratique moyenne (*MSE*) des images reconstruites gauche



(a) *art*



(b) *books*



(c) *cloth3*



(d) *cones*



(e) *midd1*



(f) *midd2*



(g) *moebius*



(h) *teddy*

FIGURE 3.7 – Les huit images stéréoscopiques utilisées dans l'expérience. Seulement les images gauche de la paire sont affichées ici.

et droite comme suit :

$$PSNR = 10 \cdot \log_{10} \left(\frac{255^2}{(MSE_l + MSE_r)/2} \right) \quad (3.14)$$

avec $MSE_k = \frac{1}{NM} \sum_{i=1}^N \sum_{j=1}^M (I_k(i, j) - \hat{I}_k(i, j))^2$, $k \in \{l, r\}$.

où (i, j) sont les coordonnées des pixels et (N, M) sont les dimensions des images. I et \hat{I} représentent les images originales et reconstruites (version décodée), respectivement.

Les performances de codage sont illustrées par les courbes de débit-distorsion (R-D) pour les différentes images, comme illustré dans la figure 3.8. L'axe vertical représente le PSNR moyen et l'axe horizontal indique le débit nécessaire pour coder les images stéréoscopiques. En plus, les gains de débit calculés selon la mesure *Bjontegaard* [128], qui est largement utilisée, sont indiqués dans le tableau 3.1.

Comme le montre la figure 3.8, la méthode *CS* est celle qui nécessite le plus grand débit, ce qui en résulte une faible performance du fait que la paire stéréo est codée sans aucun type d'asymétrie. En revanche, la méthode proposée atteint une qualité similaire tout en diminuant le débit requis. Les performances des autres méthodes de codage asymétrique dépendent du contenu des images, elles varient d'une paire stéréo à l'autre. Dans l'ensemble, la méthode *AC_Down* obtient de meilleures performances que les deux méthodes de codage asymétrique restantes (*AC_Blur* et *AC_FQuality*), mais comparée à la méthode proposée, il est clair que notre méthode est plus efficace et surpasse la méthode *AC_Down*. Pour la méthode *AC_Blur*, malgré le gain de débit obtenu, c'est celle qui offre la plus faible performance, puisque cette méthode a obtenu les plus faibles valeurs de PSNR.

La méthode proposée offre un gain de débit significatif sans la réduction de la qualité visuelle des images stéréoscopiques reconstruites. Cela a été atteint grâce à l'adaptation au contenu de la paire stéréo et aussi grâce à la prise en compte de la qualité de l'autre vue. Contrairement à la méthode *AC_FQuality*, où l'écart de qualité entre les vues est fixé d'avance, ce qui ne permet aucune adaptation.

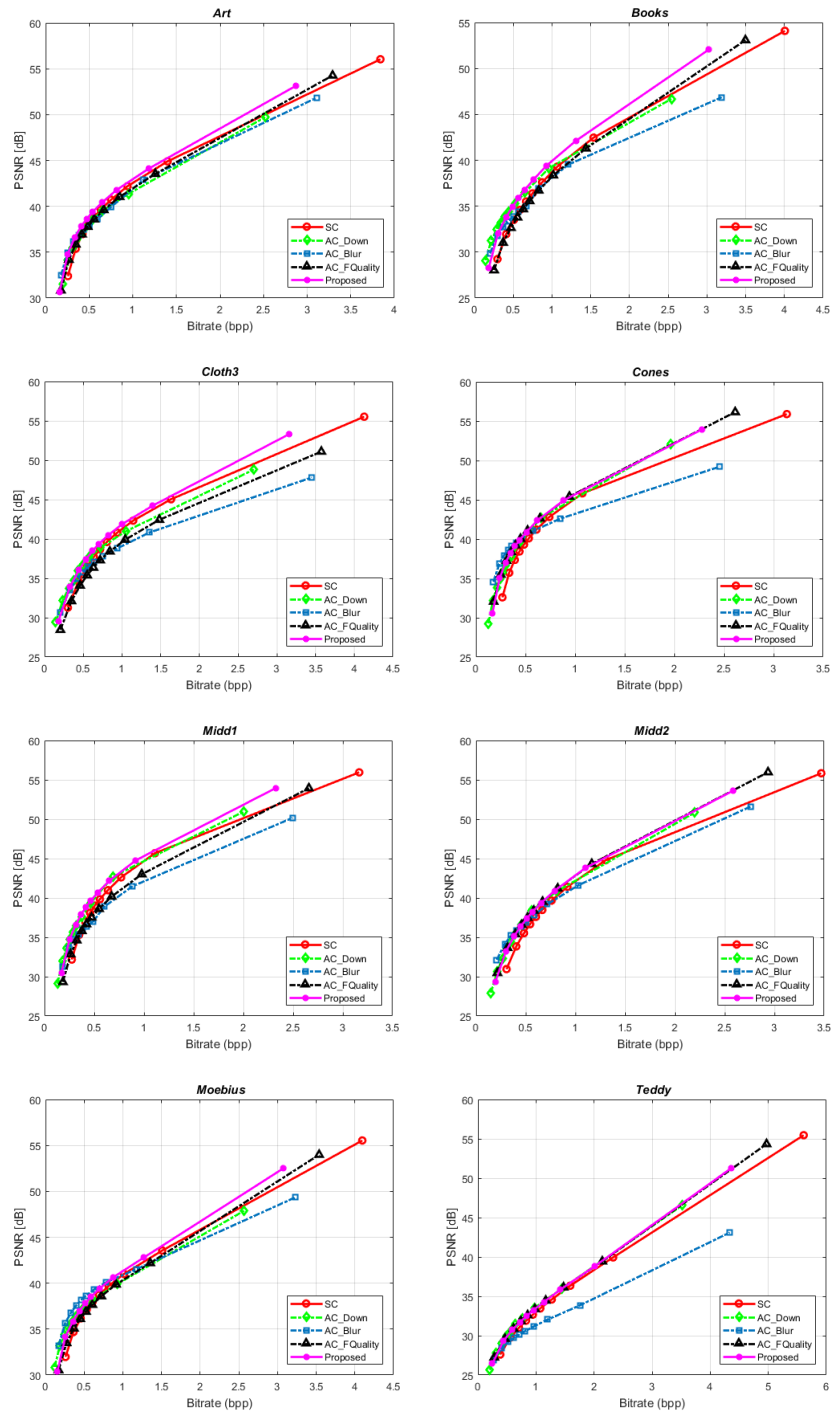


FIGURE 3.8 – Les courbes de débit-distorsion pour les huit paires d’images stéréoscopiques utilisées.

TABLE 3.1 – Les gains de bitrate et de PSNR (dB) calculés avec la mesure de Bjøntegaard. Notant que les valeurs négatives correspondent aux gains de bitrate par rapport à la méthode de codage symétrique, tandis que pour les PSNR, les valeurs négatives indiquent une réduction du PSNR.

Images	AC_Down		AC_Blur		AC_FQuality		Proposed	
	Bitrate saving	Δ PSNR	Bitrate saving	Δ PSNR	Bitrate saving	Δ PSNR	Bitrate saving	Δ PSNR
<i>Art</i>	-13.06%	0.56	-8.40%	0.61	-6.16%	0.74	-14.66%	1.07
<i>Books</i>	-11.20%	0.94	-3.32%	0.05	-2.66%	0.12	-15.32%	1.23
<i>Cloth3</i>	-5.64%	0.84	1.85%	0.03	1.20%	-0.77	-8.88%	0.98
<i>Cones</i>	-15.31%	1.09	-12.42%	0.30	-12.49%	0.89	-15.90%	1.55
<i>Midd1</i>	-13.35%	1.03	0.97%	0.14	-6.61%	0.20	-15.25%	1.13
<i>Midd2</i>	-12.33%	1.12	-2.53%	0.43	-12.01%	1.56	-12.65%	1.51
<i>Moebius</i>	-10.19%	1.15	-12.04%	0.76	-1.30%	0.35	-16.32%	1.35
<i>Teddy</i>	-8.02%	0.92	-2.93%	-0.51	-7.35%	0.87	-8.14%	0.92
Average	-11.13%	0.95	-4.85%	0.22	-5.92%	0.49	-13.39%	1.21

Les gains de débit binaire et les différences de PSNR sont donnés dans le tableau 3.1. D'après les résultats de ce tableau, le gain de débit obtenu par la méthode proposée varie de 8,14% à 16,32% avec une moyenne de 13,39% par rapport à la méthode *CS*. En plus, *AC_Down* offre un gain de débit allant de 5,64% à 15,31%, montrant ainsi de faibles performances par rapport à la méthode proposée.

Enfin, les performances les plus faibles ont été obtenues par les méthodes *AC_Blur* et *AC_Quality*, qui ont respectivement réalisé des gains de 4,85% et de 5,92% en moyenne. De plus, les valeurs de PSNR obtenues par la méthode proposée sont beaucoup plus élevées que les autres méthodes. Ajoutons à cela, pour les méthodes *AC_Blur* et *AC_Quality* dans certains cas, comme pour les paires stéréo *Cloth3* et *Teddy*, on peut noter une forte dégradation du PSNR. Enfin, ces résultats confirment clairement les conclusions précédentes selon lesquelles la méthode proposée donne de meilleurs résultats.

Selon les résultats obtenus qui sont illustrés à la figure 3.8 et au Tableau 3.1, grâce à l'ajustement automatique et efficace de l'écart de qualité entre les deux images de la paire, la méthode proposée offre un gain de débit significatif par rapport à l'approche conventionnelle, et cela sans introduire de distorsion notable dans la visualisation 3D.

3.6 Conclusion

Afin d'améliorer le codage asymétrique des images stéréoscopiques, dans ce chapitre, nous avons proposé un nouveau modèle qui nous permet d'ajuster de manière automatique et adaptative l'écart de qualité entre les images de la paire stéréo. Nous avons établi un seuil de différence en terme de qualité en modélisant la relation entre la qualité des deux vues et leur distorsion inter-vues et, également, grâce à l'inclusion du modèle inspiré du SVH qui est le BJND.

Sur la base de ce seuillage automatique de la qualité, la paire stéréo est codée de manière asymétrique, réduisant ainsi le débit nécessaire sans générer d'artefacts visibles dans la visualisation 3D. Les résultats expérimentaux montrent l'efficacité de l'approche proposée par rapport aux méthodes de codage asymétrique classiques basées sur un seuil fixe d'asymétrie.

Dans les futurs travaux, la méthode proposée peut être étendue aux cas des vidéos stéréoscopiques/multi-vues, ce qui constitue de nouveaux défis, tels que le maintien de la cohérence dans la qualité entre les images successives.

Le codage des images stéréoscopiques basé sur la colorisation

Sommaire

4.1	Introduction	68
4.2	Les méthodes de codage basées sur la colorisation	70
4.3	L'approche proposée	73
4.4	Les résultats expérimentaux	76
4.5	Conclusion	78

4.1 Introduction

Des efforts de recherche considérables ont été déployés pour développer des techniques de codage efficaces pour les images stéréoscopiques. Malgré une riche littérature et des contributions considérables [81], la plupart des méthodes proposées pour le codage d'images stéréoscopiques ont uniquement pris en compte les informations de luminance, tandis que les informations de chrominance ont reçu moins d'attention.

Peu de travaux ont été consacrés à l'amélioration de la compression de la chrominance dans les images stéréoscopiques. À l'exception de la méthode récemment proposée dans [130], où les auteurs s'appuient sur les redondances des canaux intercouleurs en plus des redondances intra- et inter-vues. La plupart des méthodes

appliquent une correction de couleur entre les images de la paire pour compenser la variation de couleur affectant les performances de compression [129].

Par ailleurs, les méthodes de codage à base de colorisation pour l'image monoscopique ont démontré leur efficacité pour réduire le débit binaire. L'idée principale derrière l'approche de codage basée sur la colorisation consiste à extraire un petit ensemble de pixels représentatifs (PR) de l'image couleur originale au niveau du codeur, où les informations de chrominance sont conservées uniquement pour ces PR. Les positions et les valeurs de chrominance sont transmises au décodeur uniquement pour le PR avec l'ensemble de la composante de luminance. Au niveau du décodeur, les valeurs de chrominance des pixels restants sont restaurées en utilisant une méthode de colorisation. Les informations de chrominance étant représentées par une très petite quantité d'informations, cela permet de réduire considérablement la bande passante de transmission. Malgré le potentiel de cette approche, elle n'a pas été explorée pour le codage d'image stéréo. De plus, on peut s'attendre à ce que l'application d'une telle approche à des images stéréoscopiques puisse fournir une réduction significative du débit, tout en maintenant une expérience visuelle 3D optimale. On peut donc supposer que, même si la reconstruction de la couleur fournit des couleurs légèrement différentes, la théorie de la suppression binoculaire [101,141] rend cette différence indiscernable dans la visualisation 3D.

Par conséquent, dans ce chapitre, nous proposons une méthode de codage d'image stéréoscopique en couleur basée sur la colorisation [144]. Dans le schéma proposé, une vue de la paire stéréo est compressée en utilisant une méthode de codage standard, alors que pour l'autre vue seule la composante de luminance est considérée pour la compression, permettant ainsi un gain de débit. L'information de chrominance de cette dernière vue est presque totalement ignorée, cela signifie qu'elle n'est pas codée et transmise, à l'exception de quelques pixels appartenant aux régions occultées qui sont considérées comme PR, et pour lesquelles les informations de couleur sont conservées. Les informations sur cet ensemble de PR sont codées efficacement en tant qu'information supplémentaire. Enfin, au décodeur, la couleur des pixels des régions non occultées est définie. Ceci est réalisé sur la base d'une technique de mise en correspondance entre les deux vues. En d'autres termes, pour ces pixels, nous essayons de transférer les informations de couleur de l'image de référence (c'est-à-dire, l'image en couleur) à l'image cible. Sur la base des informations de couleur définies pour les quelques pixels des régions occultées et non occultées en plus des informations de luminance décodées, l'algorithme de colorisation reconstruit les informations de couleur de tous les pixels restants de l'image.

Nos contributions principales consistent en l'extension du concept de codage basé sur la colorisation au cas des images stéréoscopiques. En plus, nous proposons dans ce travail une nouvelle méthode pour extraire automatiquement les PRs, et ceci en se basant sur une technique de squelettisation.

Ce chapitre est organisé comme suit. Tout d'abord, dans la section 4.2 nous présentons les travaux antérieurs sur le codage basé sur la colorisation des images monoscopiques. Par la suite nous décrivons en détail chaque étape de la méthode proposée dans la section 4.3. Les résultats expérimentaux sont rapportés et discutés dans la section 4.4. Enfin, la section 4.5 conclut ce chapitre et donne des orientations futures.

4.2 Les méthodes de codage basées sur la colorisation

Comme mentionné précédemment, les méthodes de codage à base de colorisation sont une approche de compression d'image reposant sur une technique de colorisation. La colorisation est le processus qui introduit la couleur dans une image en niveaux de gris en utilisant quelques informations de couleur fournies à travers quelques pixels représentatifs (PRs). Ces derniers sont considérés comme des graines de couleur indiquant les couleurs qui se propagent automatiquement dans les régions correspondantes de l'image, produisant ainsi une image entièrement colorisée.

Plusieurs méthodes de colorisation ont été proposées dans la littérature [142] et l'une des études représentatives et pionnières dans ce domaine est celle proposée par Levin *et al.* [126] dans laquelle la procédure de colorisation est formulée comme un problème d'optimisation. Étant donné que les informations de chrominance peuvent être représentées et reconstruites en utilisant un petit nombre de PRs, cela a ouvert la voie à une nouvelle stratégie de codage connue sous le nom de *codage basé sur la colorisation*.

Suivant cette idée, plusieurs approches ont été proposées. Dans [131–140] l'idée commune derrière la plupart des méthodes existantes est de sélectionner les PRs de l'image couleur originale au niveau du codeur, les positions et les valeurs de chrominance de ces dernières sont transmises au décodeur, tandis que le canal de luminance est compressé en utilisant des techniques de compression d'image classiques et également envoyé au décodeur. Ensuite, au niveau du décodeur, sur la base de l'information des PRs et de la luminance décodée, les informations de couleur

des pixels restants sont restaurées en utilisant une technique de colorisation.

Par conséquent, le problème principal dans la méthode de codage basée sur la colorisation est d'extraire efficacement les PRs de sorte que l'information associée soit réduite au maximum, tandis que la qualité visuelle de l'image couleur reconstruite soit difficilement compromise. Plusieurs méthodes ont été proposées pour atteindre cet objectif. Par exemple, Cheng *et al.* [131] ont proposé d'effectuer une sur-segmentation de l'image en couleur et de sélectionner de façon aléatoire les PRs à partir des régions obtenues. En suite, en utilisant l'algorithme de Levin *et al.* [126], les auteurs ont proposé un modèle d'apprentissage automatique qui prédit la couleur des pixels restants.

Cependant, l'ensemble initial de PR peut contenir une certaine redondance, comme dans les régions lisses, ce qui ne fournit pas un débit binaire optimal. Pour résoudre ce problème, Ono *et al.* [132] ont proposé de considérer les corrélations locales entre les composantes de luminance et de chrominance. Ils segmentent l'image originale en blocs et considèrent le pixel central de chaque bloc comme PR initial. Ensuite, la redondance dans l'ensemble PR initial est réduite en utilisant leurs positions et leur composante de luminance. Miyata *et al.* [133] ont proposé d'extraire les PRs sous la forme d'un ensemble de segments de ligne, qui sont extraits de manière itérative en comparant l'image en couleur reconstruite avec l'image originale.

Horiuchi *et al.* [135] ont évalué trois stratégies pour extraire les PRs : une sélection aléatoire, une sélection à partir de l'histogramme de luminance et une sélection à partir du centre de la boîte à haute densité de pixels dans l'espace CIELAB. Les résultats ont montré que cette dernière stratégie fournit de meilleurs résultats. Lee *et al.* [136] ont formulé le problème de sélection des PRs comme un problème d'optimisation. Tout d'abord, les auteurs construisent la matrice de colorisation en effectuant une segmentation à plusieurs échelles sur la base de la luminance. Ensuite, en utilisant cette matrice et les valeurs de chrominance d'origine, les PRs sont extraits en résolvant un problème de minimisation L1.

Dans la même direction et dans le but d'améliorer l'efficacité de la compression, Mishiba *et al.* [137] ont proposé une structure de matrice de colorisation basée sur la corrélation luminance-chrominance dans une zone locale en plus d'un processus de filtrage de lissage préservant les contours. Oh *et al.* [138, 139] ont proposé une méthode de régression locale itérative pour l'extraction de PR. Premièrement, les régions locales sont définies en utilisant une segmentation d'image hiérarchique, puis les paramètres de régression locale sont calculés de manière itérative pour s'adapter aux pixels majeurs dans une région locale. Afin de réduire le temps de

calcul, Uruma *et al.* [140] ont proposé d'extraire les PRs d'un ensemble d'images à résolution multiple obtenues par sous-échantillonnage. La colorisation est effectuée sur une image de résolution inférieure et propagée au niveau suivant jusqu'au niveau de haute résolution.

Presque toutes les méthodes décrites précédemment proposent d'extraire les PRs en utilisant une approche itérative ou formulée comme un problème d'optimisation, où la qualité de l'image couleur reconstruite est utilisée comme base pour décider si plus de PRs doivent être extraits ou pas. De plus, certains travaux proposent une étape supplémentaire de réduction de la redondance appliquée à l'ensemble PR initial. Tout ceci contribue à augmenter la complexité sans garantir qu'à la fin l'ensemble des PRs ne contiennent pas de redondance ou que tous les pixels requis aient été sélectionnés.

Dans notre approche, le concept de codage basé sur la colorisation est efficacement étendu au cas des images stéréoscopiques. Étant donné que les images de la paire stéréo contiennent un contenu presque similaire, puisque les deux images correspondent à la même scène à partir de différents points de vue. Ceci peut être exploité pour sélectionner les PRs non seulement à partir de l'image elle-même, comme cela est fait dans l'image monoscopique, mais aussi de l'autre image. Cela signifie que pour une seule vue de la paire stéréo, les informations de chrominance sont conservées, tandis que pour l'autre vue, seules les informations de luminance sont prises en compte pour la compression. L'utilisation de PRs à partir de l'autre vue permet au décodeur d'effectuer une colorisation par transfert de couleur entre les deux images en utilisant la mise en correspondance et sans envoyer d'informations sur les PRs.

Cependant, puisque les deux images sont acquises à partir de différentes perspectives, certaines régions peuvent apparaître dans une image mais pas dans l'autre (régions occultées). Il est évident que pour ces derniers, nous ne pouvons pas effectuer de transfert de couleur à partir de l'autre image. Par conséquent, pour ces régions occultées, les PRs doivent être extraits de l'image elle-même. Pour y parvenir, nous proposons de l'effectuer en utilisant la squelettisation, et chaque squelette pour chacune de ces régions est considéré comme PR.

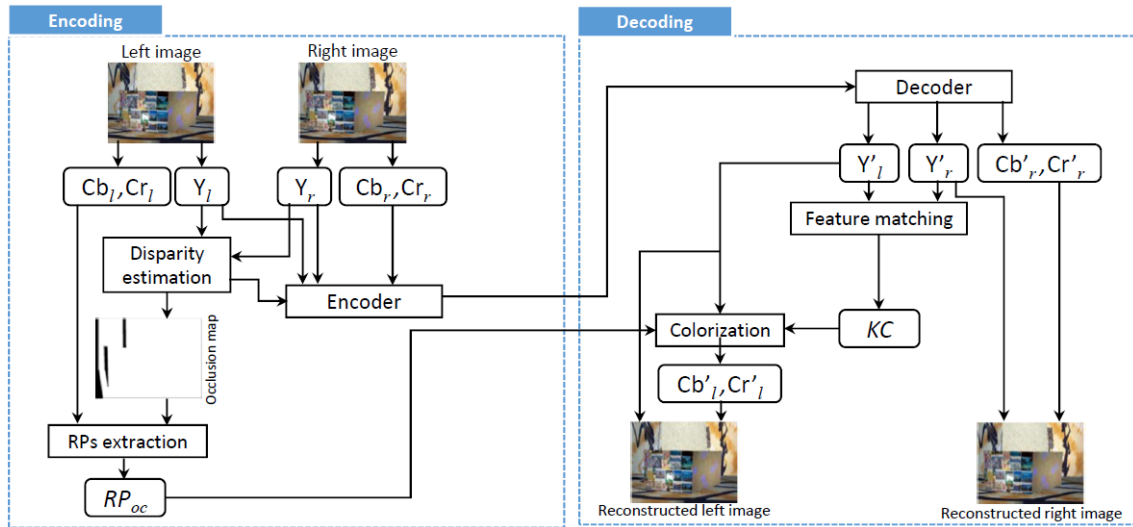


FIGURE 4.1 – Le schéma proposé pour le codage d'image stéréoscopique basé sur la colorisation, de l'encodage au décodage.

4.3 L'approche proposée

Le schéma générale de la méthode proposée est illustré dans la figure 4.1. Dans la suite de cette section, nous décrivons en détail chaque composante de ce schéma. Étant donné une paire stéréo composée de deux images, gauche I_l et droite I_r , nous considérons I_r comme l'image de référence et I_l comme image cible. Notre approche suppose une représentation dans l'espace colorimétrique YCbCr. Cela permet de gérer séparément la composante de luminance Y et les deux composantes de chrominance (C_b et C_r). Il est important de noter que les composantes de luminance Y_r et de chrominance (C_{b_r} et C_{r_r}) de I_r sont codées à l'aide d'un codeur d'images fixes existant, tandis que pour I_l , seule la composante de luminance est codée en utilisant une prédiction de I_r et cela en utilisant la compensation de disparité. Cependant, pour les informations de chrominance de I_l , une étape de prétraitement est effectuée pour ne conserver qu'un petit nombre de PRs, avant d'être envoyée au codeur.

Pour y parvenir, la carte de disparité est d'abord calculée en utilisant l'algorithme de mise en correspondance décrit dans [120]. Sur la base des informations de disparité calculées, une classification en régions occultées et non occultées est effectuée. Par conséquent, chaque pixel de la vue de gauche doit être caractérisé comme appartenant à une classe occultée ou non occultée. Pour ce faire, nous appliquons un test de cohérence entre les pixels correspondants des deux images gauche et droite. Cette approche est très populaire dans l'appariement stéréoscopique pour identifier

les pixels occultés dans les cartes de disparités [121].

Ce test exploite les deux cartes de disparité : la carte de disparité de gauche à droite d_l , et la carte de disparité de droite à gauche d_r . Les pixels correspondants sont considérés comme occultés si la somme de leurs valeurs de disparité est supérieure à un seuil T_{LR} (nous utilisons $T_{LR} = 1$) :

$$|d_l(i, j) + d_r(i, j - d_l(i, j))| \geq T_{LR} \quad (4.1)$$

où $d_k(i, j)$ est la valeur de disparité du pixel à la position (i, j) dans l'image k .

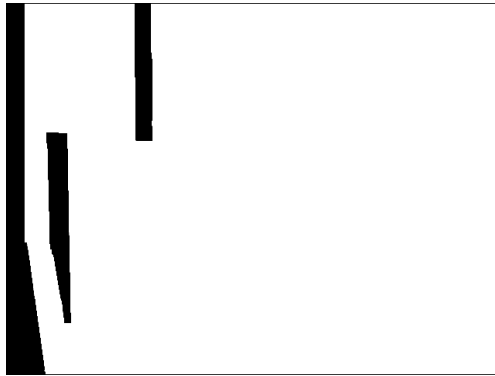
Une fois la classification effectuée, les informations de couleur des pixels appartenant à des régions non occultées sont complètement ignorées par le codeur. Parce que, ces derniers sont restaurées par colorisation (par transfert de couleur) en utilisant les pixels correspondants de I_r . En revanche, étant donné que les pixels appartenant à des régions occultées ne sont visibles que dans une vue et n'ont pas de pixels correspondants dans I_r , cela rend le transfert de couleur à partir de I_r au niveau du décodeur impossible. Par conséquent, pour les régions occultées, nous extrayons un petit ensemble de PRs en utilisant une nouvelle approche.

Pour extraire les PRs, différentes méthodes ont été proposées dans la littérature [133, 136]. Dans notre approche, nous proposons de l'effectuer en utilisant la squelettisation [122]. Afin d'extraire les PRs des régions occultées, nous estimons un squelette pour chacune de ces régions en utilisant la méthode proposée dans [123], comme illustré à la figure 4.2. Ainsi, les pixels appartenant aux squelettes sont considérés comme des PRs de ces régions et ils sont désignés par PR_{oc} . Les valeurs de position et de chrominance de ce petit nombre de PR sont envoyées au décodeur. Ces derniers sont codés en tant que vecteur à quatre dimensions, c'est-à-dire deux pour les coordonnées spatiales et deux pour les valeurs C_b et C_r . Ainsi, si nous assignons 2 octets pour chaque élément du vecteur, cela donne 8 octets pour chaque PR. Cependant, pour permettre une visualisation 3D confortable, la proportion de régions occultées est généralement très faible par rapport au reste de l'image. Ainsi, l'augmentation du débit binaire liée à cette information est très faible.

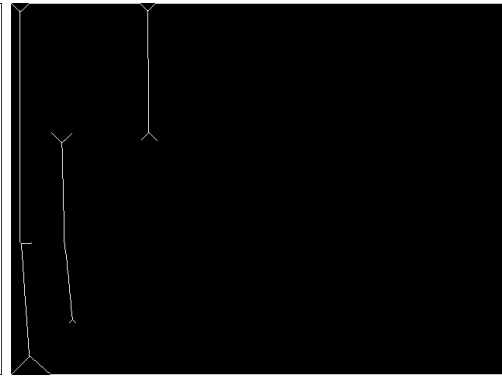
Au niveau du décodeur, une fois que l'ensemble PR_{oc} reçu et I_r décodé (noté I'_r), ainsi que la composante de luminance de I_l (notée Y'_l), ils sont utilisés pour récupérer les composantes de couleur de certains pixels non-occultés de I_l . Ceci est réalisé à travers la mise en correspondances éparses entre les deux images. Nous avons choisi la méthode SIFT (scale invariant feature transform) [124] pour trouver l'ensemble des correspondances, car elle est considérée comme l'une des techniques



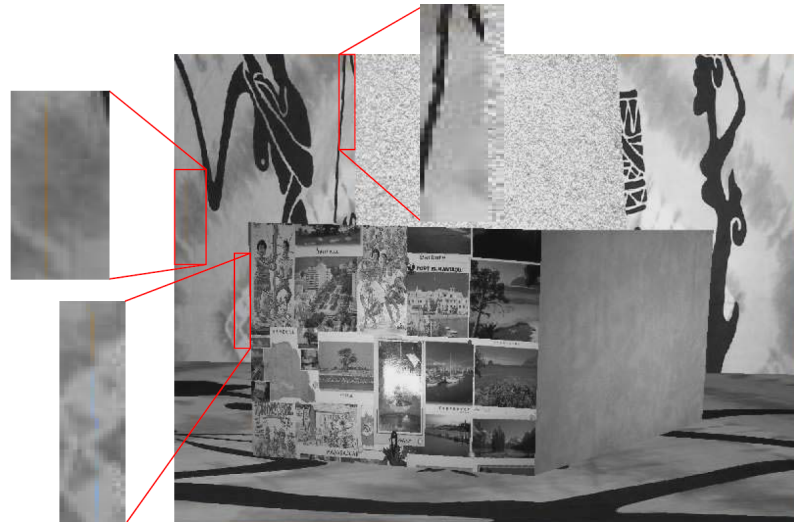
(a) Image couleur gauche



(b) Carte d'occlusion



(c) Résultat de la squelettisation



(d) Image gauche en niveau gris avec squelette

FIGURE 4.2 – Exemple de squelette utilisé pour extraire les PRs.

de correspondance d'images les plus efficaces.

La méthode SIFT est appliquée à Y_l' , et Y_r' , conduisant à deux ensembles de points-clés SIFT. Ensuite, les descripteurs des points clés des deux ensembles sont extraits et comparés. Afin d'éliminer les correspondances non fiables, l'ensemble des points-clés correspondants est affiné par un filtrage géométrique à travers l'algorithme RANSAC (RANdom SAmple Consensus) [125].

Sur la base de ces correspondances fiables, le transfert de couleur est établi entre les points-clés correspondants. En d'autres termes, nous attribuons à chaque point-clé de Y_l' la même valeur de couleur que son point-clé correspondant de Y_r' . Cet ensemble de points-clés appariés de Y_l' pour lesquels la couleur a été définie est noté KC.

Enfin, l'ensemble des PR_{oc} déjà défini, en plus de l'ensemble des pixels colorés KC ainsi que le Y_l' décodé, sont utilisés comme entrée pour la méthode de colorisation. Pour ce dernier, nous avons opté pour le très répandu algorithme de Levin *et al.* [126], cet algorithme de colorisation est utilisé pour reconstruire les informations de couleur de tous les pixels restants de Y_l' . En utilisant les quelques pixels pour lesquels les informations de couleur sont déjà définies (RP_{oc} et KC) avec les informations de luminance (Y_l'), l'algorithme de colorisation propage la couleur définie au reste de l'image en supposant que les pixels voisins avec une luminance similaire devraient également acquérir une chrominance similaire.

De cette manière, nous pouvons obtenir une réduction du débit binaire, car une partie considérable des composants de chrominance n'a pas été envoyée. En outre, cette méthode permet d'éviter la variation des couleurs entre les images, comme cela est généralement fait avec les méthodes de correction des couleurs [129].

4.4 Les résultats expérimentaux

Afin d'évaluer les performances de l'approche proposée, nous avons utilisé huit paires d'images stéréo obtenues à partir de la base de données des images stéréoscopiques *Middlebury* [127]. La méthode proposée est comparée à la méthode de codage classique qui code à la fois les informations de luminance et de chrominance des deux images sans aucune astuce (appelée par la suite méthode de codage d'ancrage).

Pour les deux méthodes comparées, nous avons utilisé JPEG 2000 [118] comme standard de codage. Cependant, il est important de noter que nous pouvons utiliser

n'importe quel autre encodeur d'image fixe de l'état de l'art, car notre approche peut être adaptée facilement. Comme mentionné précédemment, l'image de droite est codée comme image de référence et la gauche comme image cible. La dernière image est prédite à partir de l'image de droite en utilisant la prédiction inter-image. Ainsi, la disparité est codée sans perte à l'aide d'un codeur arithmétique, tandis que les images droite et résiduelle sont codées en JPEG 2000 pour des débits à trois niveaux (bas, moyen et haut) correspondant aux quatre points cibles $\{0.15, 0.2, 0.3, 0.4\}$, $\{0.5, 0.6, 0.7, 0.8\}$ et $\{0.9, 1, 1.1, 1.2\}$ bpp, respectivement. Les résultats sont présentés sous forme de comparaison quantitative en termes de débit moyen R_{av} et de mesure de PSNR correspondante.

Puisque notre approche implique l'envoi d'une petite quantité d'information liée à PR_{oc} , elle est donc prise en compte lors du calcul du débit requis comme suit :

$$R_{av} = \frac{R_r + R_{res} + R_d}{2} + R_{PR_{oc}} \quad (4.2)$$

$$PSNR = 10. \log_{10} \left(\frac{255^2}{(MSE_l + MSE_r)/2} \right) \quad (4.3)$$

où R_d , R_r et R_{res} désignent respectivement le débit de la carte de disparité et des images droite et résiduelles. $R_{PR_{oc}}$ représente le débit binaire lié à PR_{oc} . Ainsi, pour la méthode de codage d'ancrage, PR_{oc} est fixé à 0. MSE_l et MSE_r sont les erreurs quadratiques moyennes des images gauche et droite reconstruites, respectivement. Il est important de noter que, pour notre méthode, le calcul du PSNR est effectué après le processus de colorisation.

Le tableau 4.1 présente le gain de débit binaire et le gain PSNR moyen calculés en utilisant la mesure Bjøntegaard [128]. Il fournit également les valeurs moyennes pour l'ensemble des image de test. Ce tableau montre que le gain de débit obtenue par la méthode proposée est compris environ entre 13% et 20% par rapport à la méthode de codage conventionnelle. De même, en termes de PSNR moyen, les gains sont d'environ de 1.28 – 2 dB.

Le débit binaire lié à PR_{oc} (c'est-à-dire, $R_{PR_{oc}}$) dépend de la proportion des régions occultées dans l'image. Cependant, la taille de ces régions n'excède généralement pas 30% et se situe généralement entre 5% et 15%. Malgré le coût de cette information secondaire, qui est faible, la méthode proposée surpasse toujours la méthode de codage conventionnelle, comme l'illustrent les résultats fournis.

La figure 4.3 montre les courbes de débit-distorsion pour l'image gauche de certaines paires stéréo. Selon les résultats obtenus, grâce à la gestion efficace des

TABLE 4.1 – La comparaison des gains de bitrate et des moyennes de PSNR (dB), mesurés par rapport à la méthode de codage conventionnelle en utilisant la mesure de Bjøntegaard [128].

Image	Δ Bitrate saving (%)			Δ PSNR gain (dB)		
	low	middle	high	low	middle	high
<i>Art</i>	-20.52	-18.37	-15.56	0.86	0.94	1.30
<i>Laundry</i>	-18.48	-13.64	-15.24	1.95	2.28	2.34
<i>Reindeer</i>	-13.15	-12.43	-10.72	1.38	1.81	2.06
<i>Cones</i>	-15.92	-12.98	-13.07	0.94	0.96	1.28
Average	-14.84	-12.42	-12.06	1.12	1.25	1.53

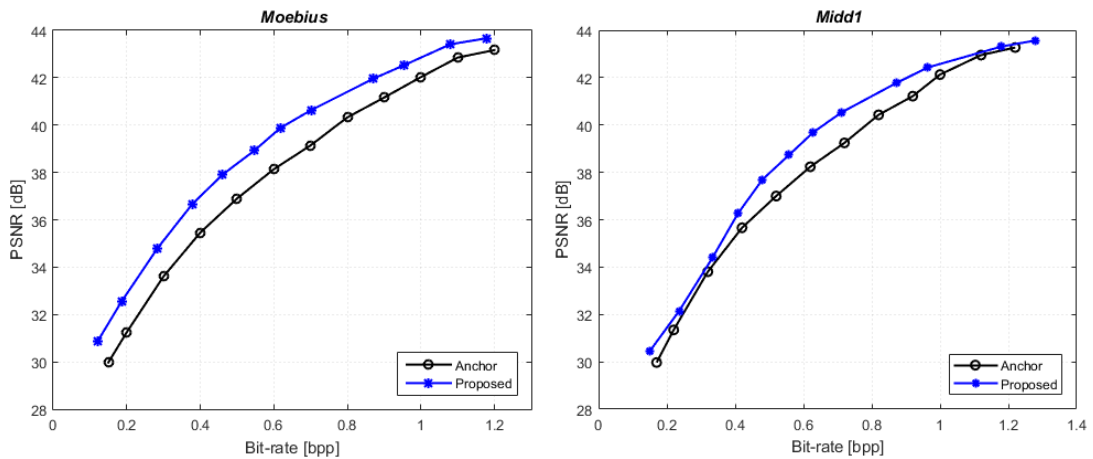


FIGURE 4.3 – Courbes de débit-distorsion pour l’image gauche des paires stéréo *Moebius* et *Midd1*.

composants de chrominance, la méthode proposée permet un gain de débit significatif par rapport à l’approche de codage conventionnelle.

En plus de ces résultats quantitatifs, nous présentons dans la figure 4.4 un exemple d’images gauches décodées en utilisant la méthode de codage conventionnelle et notre approche. Cette figure illustre clairement que notre méthode fournit des résultats fiables par rapport à la méthode de codage conventionnelle ancrage.

4.5 Conclusion

Dans ce chapitre, une nouvelle méthode de codage stéréoscopique basée sur la colorisation a été proposée. Notre approche consiste à coder une vue de la paire stéréo avec un codeur standard, tandis que pour l’autre vue seule la composante de

luminance est codée, l'information de chrominance est presque totalement ignorée, à l'exception de quelques pixels des régions occultées qui sont considérées comme des PRs, et pour lesquelles les informations de couleur sont conservées.

Afin de réduire significativement le débit binaire, nous avons aussi proposé une nouvelle méthode d'extraction des PRs basée sur la technique de squelettisation. De plus, puisque côté décodeur, ces informations de couleur sont transférées à partir de l'autre image, cela apporte un autre avantage en ce qui concerne la calibration des couleurs entre les images, ce qui améliore l'expérience de visualisation 3D.

Les résultats d'évaluation ont montré que la méthode proposée permet de réaliser des économies de débit significatives sans perte de qualité visuelle. Par conséquent, nous pensons que l'extension de cette méthode aux cas des vidéos stéréoscopiques/multi-vues présente de grands potentiels, et constitue de nouveaux défis tels que le maintien de la couleur constante entre les trames successives.



FIGURE 4.4 – Exemples d’images gauche reconstruites. Image cible originale (à gauche), méthode de codage conventionnelle (au milieu) et notre méthode (à droite).

Conclusion et perspectives

La compression des images stéréoscopiques constitue une brique fondamentale dans la chaîne de diffusion des médias S-3D. Cette thèse a été principalement consacrée à la conception et au développement de nouveaux schémas de codage efficaces pour les images stéréoscopiques. Dans ce qui suit, nous résumons les principales contributions de cette thèse et nous mentionnons quelques perspectives possibles à la fin de chaque paragraphe.

La première contribution de ce travail est basée sur la théorie de suppression binoculaire, et plus précisément concerne le développement d'une méthode de codage asymétrique. Dans cette méthode, une image de la paire stéréo peut être codée avec une qualité inférieure par rapport à l'autre. Pour réaliser cela nous avons proposé un nouveau modèle ajustant de manière automatique et adaptative l'écart de qualité entre les images de la paire stéréoscopique. Nous avons établi un seuil de qualité en modélisant la relation entre la qualité des deux images et leur distorsion inter-vues et, également, grâce à l'inclusion du modèle inspiré du SVH. Sur la base de ce seuillage automatique de la qualité, la paire stéréo est codée de manière asymétrique, réduisant ainsi le débit nécessaire sans induire d'artefacts visibles dans la visualisation 3D.

Les résultats expérimentaux montrent l'efficacité de l'approche de codage proposée par rapport aux méthodes de codage asymétrique classiques basées sur un seuil fixe d'asymétrie. Dans les travaux futurs, la méthode proposée peut être étendue aux cas des vidéos stéréoscopiques/multi-vues.

La deuxième contribution de ce travail concerne le codage de l'information de chrominance des images stéréoscopiques. Pour atteindre cet objectif, nous avons proposé d'étendre le concept de la méthode de codage basée sur la colorisation, qui

a été largement explorée pour les images 2D, au cas des images stéréoscopiques. Les informations chromatiques des images codées sont transmises au décodeur uniquement pour quelques pixels représentatifs. Pour extraire automatiquement ces pixels, nous avons proposé une nouvelle méthode basée sur la squelettisation. En raison du fait que seulement une petite information de couleur a été envoyée au décodeur, ceci permet de réduire significativement le débit binaire requis pour délivrer la paire stéréoscopique. De plus, puisque les informations de couleur sont transférées à partir de l'autre image, cela apporte un autre avantage en ce qui concerne la calibration des couleurs entre les vues, ce qui améliore l'expérience de la visualisation 3D.

Les résultats expérimentaux montrent l'efficacité de l'approche de codage proposée par rapport à l'approche conventionnelle. Les directions futures de ce travail peuvent être l'amélioration du codage du squelette, l'extension de la méthode proposée à la vidéo stéréoscopique et la validation des résultats par des expériences subjectives.

Bibliographie

- [1] Michel. Benard, “ Restitution automatique en stéréophotogrammétrie”, *Thèse ENST Département ISSV laboratoire Image*, 1983. (Cité à la page 8.)
- [2] Nicholas. Ayache, “Vision stéréoscopique et perception multisensorielle”, *Inter éditions science informatique*, 1989. (Cité à la page 8.)
- [3] Jean. Carré, (1971). “Lecture et exploitation des photographies aériennes”, *Editions EYROLLES Paris*, vol. 2,. Automatique, INRIA, juillet, 1996. (Cité à la page 8.)
- [4] Z. Zhang, G. Xu, “A general expression of the fundamental matrix for both perspective and affine cameras”, *In International Joint Conference on Artificial Intelligence*, vol. 2, pp 1502-1507, Nagoya, Japan, 1997. (Cité à la page 13.)
- [5] Z. Zhang, “Determining the epipolar geometry and its uncertainty : a review”. *International Journal on Computer Vision*, vol. 27, no. 2, pp. 161–195, Mars 1998. (Cité à la page 13.)
- [6] X. Armangué et J. Salv, “Overall view regarding fundamental matrix estimation”, *International Journal of Image and Vision Computing, IVC*, vol. 21, no. 2, pp. 205-220, février 2003. (Cité à la page 13.)
- [7] R. Hartley, R. Gupta, “ Computing matched-epipolar projections”, *In IEEE Conference on Computer Vision and Pattern Recognition*, pp. 549-555, New York, United States, 1993. (Cité à la page 14.)
- [8] A. Fusiello, E. Trucco, A. Verri, “A compact algorithm for rectification of stereo pairs”. *International Journal of Machine Vision and Applications*, vol. 12, no. 1, pp. 16-22, 2000. (Cité à la page 14.)

- [9] D. V. Papadimitriou, T. J. Dennis, “ Epipolar line estimation and rectification for stereo image pairs”. *IEEE Transactions on Image Processing*, vol. 5, no. 4, pp. 672-677, 1996. (Cité à la page 14.)
- [10] F. Devernay, “ Vision stéréoscopique et propriétés différentielles des surfaces ”. *Thèse de doctorat, Institut National Polytechnique, INP, Grenoble, France*, février 1997. (Cité à la page 14.)
- [11] F. Jurie, “ Reconnaissance d’objets volumiques par mise en correspondance d’indices visuels ”. *Traitement du signal, TS*, vol. 18, no. 5, pp. 321-345, 2001. (Cité à la page 15.)
- [12] E. G. M. Petrakis, A. Diplaros et E. Milios, “ Matching and Retrieval of Distorted and Occluded Shapes Using Dynamic Programming”. *IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI*, vol. 24, no. 11, pp. 321-345, novembre 2002. (Cité à la page 15.)
- [13] J. Matas, O. Chum, M. Urban et T. Pajdla, “ Robust Wide Baseline Stereo from Maximally Stable Extremal Regions ”. *Dans proceedings of British Machine Vision Conference, BMVC*, vol. 1, pp. 384-393, Cardiff, Royaume-Uni, septembre 2002. (Cité à la page 15.)
- [14] R. I. Hartley, P. Sturn, “ Triangulation”. *Computer Vision and Image Understanding*, vol. 68, no. 2, pp. 146-157, 1997. (Cité à la page 16.)
- [15] V. R. Dhond, J. K. Aggarwal , “ Structure from stereo - a review ”, *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 19, no. 6, pp. 1489-1510, 1989. (Cité à la page 18.)
- [16] Y. Otha, T. Kanade, “ Stereo by intra and inter-scanline search using dynamic programming”. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 7, no. 2, pp. 139-154, 1985. (Cité à la page 19.)
- [17] G. Medioni, R. Nevatia, “ Segment-based stereo matching”. *Computer Vision, Graphics and Image Processing*, vol. 31, pp. 2-18, 1985. (Cité à la page 19.)
- [18] M. Alkanhal, D. Turaga, T. Chen, “Correlation based search algorithms for motion estimation”. *In Picture Coding Symposium*, pp 99-102, Portland, 1999. (Cité à la page 20.)
- [19] B. Zitova, J. Flusser, “ Image registration methods : a survey ”. *Image and Vision Computing*, vol. 21, no. 11, pp. 977-1000, 2003. (Cité à la page 20.)

- [20] T. Kanade, M. Okutomi, "A stereo matching algorithm with an adaptive window : Theory and experiment". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, no. 9, pp. 920-932, 1994. (Cit     la page 21.)
- [21] O. Woo, A. Ortega, " Overlapped block disparity compensation with adaptive windows for stereo image coding", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 10, no. 2, pp. 3194-200, 2000. (Cit     la page 21.)
- [22] Y. Boykov, O. Veksler, R. Zabih., "Fast approximate energy minimization via graph cuts". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 11, pp. 1222-1239, 2001. (Cit     la page 21.)
- [23] M. Xing, C. Hexin, Z. Yan, "Stereo image coding method using stereo matching with difference based adaptive searching windows", *In IEEE International Workshop on Imaging Systems and Techniques*, pp. 373-376, Shenzhen, 2009. (Cit     la page 21.)
- [24] O. Veksler. "Stereo correspondence by dynamic programming on a tree". *In IEEE Conference Proceedings of Computer Vision and Pattern Recognition*, vol. 2, pp. 384-390, San Diego, United States, 2005. (Cit     la page 21.)
- [25] C. Kim, K. Lee, B. Choi, S. Lee, " A dense stereo matching using two-pass dynamic programming with generalized ground control points". *In IEEE Conference Proceedings of Computer Vision and Pattern Recognition*, vol. 2, pp. 1075-1082, San Diego, United States, 2005. (Cit     la page 21.)
- [26] V. Kolmogorov, R. Zabih, "Computing visual correspondence with occlusions using graph cuts". *In IEEE International Conference on Computer Vision*, vol. 2, pp. 508-515, Vancouver, BC, Canada, 2001. (Cit     la page 21.)
- [27] N. Slesareva, A. Bruhn, J. Weickert, "Optic flow goes stereo : A variational method for estimating discontinuity-preserving dense disparity maps". *In the 27th DAGM Symposium*, vol. 3663, pp. 33-40, Vienna, Austria, 2005. (Cit     la page 21.)
- [28] S. Kosov, T. Thormahlen, H.P. Seidel, "Accurate real-time disparity estimation with variational methods". *In International Symposium on Visual Computing*, vol. 5875, pp. 796-807, Las Vegas, United States, 2009. (Cit   aux pages 21 and 22.)
- [29] L. Alvarez, R. Deriche, J. Sanchez, J. Weickert, "Dense disparity map estimation respecting image discontinuities : A PDE and scale-space based approach". *Journal of Visual Communication and Image Representation*, vol.13, pp. 3-21, 2002. (Cit   aux pages 22 and 39.)

- [30] A. C. Bovik, "Handbook of Image and Video Processing". *Academic Press*, San Diego, CA, 2000. (Cité à la page 24.)
- [31] C. C. Cutler, "Differential Quantization of Communication Signals". *first filling*, U.S. patent number 2605361, 1950. (Cité à la page 26.)
- [32] L. M. Marvell, A. S. Khayrallah, C. G. Boncelet, "Robust source coding of images for very noisy channels". *IEEE Transactions on Signal Processing*, vol. 47, no. 4, pp.1198-1202, 1999. (Cité à la page 26.)
- [33] Y. Hua, W. Liu, "Generalized karhunen-loeve transform". *IEEE Signal Processing Letters*, vol. 5, no. 6, pp. 141-142, 1998. (Cité à la page 27.)
- [34] K. R. Rao, P. Yip, "Discrete Cosine Transform : Algorithms, Advantages, Applications". *Academic Press*, London, 1990. (Cité à la page 27.)
- [35] S. Mallat, "A wavelet tour of signal processing". *Academic Press*, San Diego, 1998. (Cité à la page 27.)
- [36] W. B. Pennebaker, J. L. Mitchell, "JPEG Still Image Data Compression Standard". *Kluwer Academic Publishers*, Norwell, MA, USA, 1993. (Cité à la page 27.)
- [37] H. S. Malvar, "Signal processing with lapped transforms". *Artech House*, Inc., 1992. (Cité à la page 28.)
- [38] D. Taubman, M. Marcellin, "JPEG2000 : Image Compression Fundamentals, Standards and Practice". *Kluwer Academic Publishers*, Norwell, MA, USA, 2001. (Cité aux pages 28, 30, 33, and 34.)
- [39] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image coding using wavelet transform". *IEEE Transactions on Image Processing*, vol. 1, no. 2, pp. 205–220, 1992. (Cité à la page 28.)
- [40] A. Gersho, R. M. Gray, "Vector quantization and signal compression". *Springer*, 1992. (Cité à la page 28.)
- [41] G. K. Wallace, "The JPEG still picture compression standard". *Communications of the ACM*, vol. 34, no. 4, pp. 30-44, 1991. (Cité aux pages 29, 30, 34, and 43.)
- [42] M. R. Nelson. "LZW data compression". *Dr. Dobb's Journal*, vol. 14, no. 10, pp. 29–36, 1989. (Cité à la page 30.)
- [43] J. E. Fowler, B. Pesquet-Popescu, "An overview on wavelets in source coding, communications, and networks". *EURASIP Journal on Image and Video Processing*, 2007. (Cité à la page 30.)

- [44] J. M. Shapiro, “Embedded image coding using zerotrees of wavelet coefficients”. *IEEE Transactions on Image Processing*, vol. 41, no. 12, pp. 3445–3462, 1993. (Cité aux pages 31 and 44.)
- [45] D. Taubman, “High performance scalable image compression with EBCOT”. *IEEE Transactions on Image Processing*, vol. 9, no. 7, pp. 1158–1170, 2000. (Cité à la page 33.)
- [46] P. Seuntjens, L. Meesters, W. Ijsselsteijn, “Perceived quality of compressed stereoscopic images : Effects of symmetric and asymmetric jpeg coding and camera separation”. *ACM Transactions on Applied Perception*, vol. 3, no. 2, pp. 95–109, 2006. (Cité aux pages 34 and 51.)
- [47] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra, “Overview of the H. 264/AVC video coding standard”. *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 560–576, 2003. (Cité aux pages 35 and 45.)
- [48] W. Woo, A. Ortega, “Optimal blockwise dependent quantization for stereo image coding”. *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 9, no. 6, pp. 861 –867, 1999. (Cité aux pages 36 and 45.)
- [49] N. V. Boulgouris, M. G. Strintzis, “A family of wavelet-based stereo image coders”. *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 12, no. 10, pp. 898–903, 2002. (Cité aux pages 36, 42, and 44.)
- [50] S. V. Leuven, J. D. Cock, G. V. Wallendael, R. V. de Walle, R. Garrido-Cantos, J. L. Martinez, and P. Cuenca. “Combining openand closed-loop architectures for H.264/AVC-to-SVC transcoding”. In *International Conference on Image Processing*, pp. 1661–1664, 2011. (Cité à la page 37.)
- [51] D. Scharstein, R. Szeliski, “A taxonomy and evaluation of dense two-frame stereo correspondence algorithms”. *International journal of computer vision*, vol. 47, no. 3, pp. 7–42, 2002. (Cité à la page 37.)
- [52] M. E. Lukacs, “Predictive coding of multi-view point image sets”. *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 521–524, 1986. (Cité aux pages 38 and 48.)
- [53] R. Darazi, A. Gouze, B. Macq, “Adaptive lifting scheme-based method for joint coding 3D-stereo images with luminance correction and optimized prediction”. In *International Conference on Acoustics, Speech and Signal Processing*, pp. 917–920, Taipei, Taiwan, 2009. (Cité aux pages 38 and 42.)

- [54] Q. Jiang, J. J. Lee, M. H. Hayes, “A wavelet based stereo image coding algorithm”. In *International Conference on Acoustics, Speech, and Signal Processing*, vol. 6, pp. 3157–3160, Phoenix, Arizona, USA, 1999. (Cit     la page 38.)
- [55] J. N. Ellinas, M. S. Sangriotis, “Stereo image compression using wavelet coefficients morphology”. *Image and Vision Computing*, vol. 22, no. 4, pp. 281–290, 2004. (Cit   aux pages 38 and 44.)
- [56] M. Kaaniche, A. Benazza-Benyahia, B. Pesquet-Popescu, J.- C. Pesquet, “Vector lifting schemes for stereo image coding”. *IEEE Transactions on Image Processing*, vol. 18, no. 11, pp. 2463–2475, 2009. (Cit   aux pages 39 and 42.)
- [57] W. Woo, A. Ortega, “Overlapped block disparity compensation with adaptive windows for stereo image coding”. *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 10, no. 2, pp. 194 –200, 2000. (Cit     la page 39.)
- [58] J. N. Ellinas, M. S. Sangriotis, “A novel stereo image coder based on quad-tree analysis and morphological representation of wavelet coefficients”. In *International Conference on Information Technology and Quality*, Athens, Greece, 2004. (Cit   aux pages 39 and 40.)
- [59] M. Kaaniche, W. Miled, B. Pesquet-Popescu, A. Benazza-Benyahia, and J.- C. Pesquet, “Dense disparity map representations for stereo image coding”. In *International Conference on Image Processing*, pp. 725–728, Cairo, Egypt, 2009. (Cit   aux pages 39 and 40.)
- [60] J. N. Ellinas and M. S. Sangriotis, “Stereo image coder based on MRF analysis for disparity estimation and morphological encoding”. In *International Symposium on 3D Data Processing, Visualization and Transmission*, pp. 852–859. IEEE, 2004. (Cit   aux pages 39 and 40.)
- [61] D. Tzovaras, N. Grammalidis, and M. G. Strintzis, “Object-based coding of stereo image sequences using joint 3-D motion/disparity compensation”. *IEEE Transactions on Circuits and systems for video technology*, vol. 7, no. 2, pp. 312 –327, 1997. (Cit   aux pages 39 and 40.)
- [62] J. Jiang, E. A. Edirisinghe, “A hybrid scheme for low bit-rate coding of stereo images”. *IEEE Transactions on Image Processing*, vol. 11, no. 2, pp. 123–134, 2002. (Cit   aux pages 39 and 40.)
- [63] L. F. R. Lucas, N. M. M. Rodrigues, E. A. B. da Silva, S. M. M. de Faria, “Stereo image coding using dynamic template-matching prediction”. In *IEEE International Conference on Computer as a Tool*, pp. 1–4, 2011. (Cit     la page 39.)

- [64] W. Woo, A. Ortega, “Stereo image compression based on disparity field segmentation”. In *SPIE Conference on Visual Communications and Image Processing*, vol. 3024, pp. 391–402, 1997. (Cité à la page 41.)
- [65] U. Ahlvers, U. Zoelzer, S. Rechmeier, “FFT-based disparity estimation for stereo image coding”. In *International Conference on Image Processing*, vol. 1, pp. 761–764, 2003. (Cité à la page 41.)
- [66] M. S. Moellenhoff, M. W. Maier, “Transform coding of stereo image residuals”. *IEEE Transactions on Image Processing*, vol. 7, no. 6, pp. 804–812, 1998. (Cité aux pages 41, 43, and 44.)
- [67] T. Frajka, K. Zeger, “Residual image coding for stereo image compression”. *Optical Engineering*, vol. 42, no. 1, pp. 182–189, 2003. (Cité à la page 42.)
- [68] A. Maalouf, M.-C. Larabi, “Bandelet-based stereo image coding”. In *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 698–701, 2010. (Cité à la page 42.)
- [69] T. Palfner, A. Mali, E. Muller, “Optimized progressive coding of stereo image using discrete wavelet transform”. In *World Multiconference on Systemics, Cybernetics and Informatics*, vol. 1, pp. 35–39, 2002. (Cité à la page 44.)
- [70] A. Said, W. A. Pearlman, “A new, fast, and efficient image codec based on set partitioning in hierarchical trees”. *IEEE Transactions on Circuits and systems for video technology*, vol. 6, no. 3, pp. 243–250, 1996. (Cité aux pages 32 and 44.)
- [71] F. Shao, G. Jiang, W. Lin, M. Yu, Q. Dai, “Joint bit allocation and rate control for coding multi-view video plus depth based 3D video”. *IEEE Transactions on MultiMedia*, vol. 15, no. 8, pp. 1843–1854, 2013. (Cité à la page 45.)
- [72] B. Rajaei, T. Maugey, H.-R. Pourreza, P. Frossard, “Rate-distortion analysis of multiview coding in a DIBR framework”. *Annals of telecommunications*, pp. 1–14, 2012. (Cité à la page 45.)
- [73] A. Gelman, P. L. Dragotti, V. Velisavljevic, “Multiview image coding using depth layers and an optimized bit allocation”. *IEEE Transactions on Image Processing*, vol. 21, no. 9, pp. 4092–4105, 2012. (Cité à la page 45.)
- [74] S.-W. Wu, A. Gersho “Rate-constrained optimal block-adaptive coding for digital tape recording of HDTV”. *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 1, no. 1, pp. 100–112, 1991. (Cité à la page 45.)

- [75] T. André, M. Cagnazzo, M. Antonini, M. Barlaud, “JPEG2000-compatible scalable scheme for wavelet-based video coding”. *EURASIP Journal on Image and Video Processing*, 2007. (Cité à la page 45.)
- [76] A. Fraysse, B. Pesquet-Popescu, J.-C. Pesquet, “On the uniform quantization of a class of sparse sources”. *IEEE Transactions on Information Theory*, vol. 55, no. 7, pp. 3243–3263, 2009. (Cité à la page 45.)
- [77] A. Gyorgy, T. Linder, K. Zeger, “On the rate-distortion function of random vectors and stationary sources with mixed distributions”. *IEEE Transactions on Information Theory*, vol. 45, no. 6, pp. 2110–2115, 1999. (Cité à la page 45.)
- [78] A. Vetro, A. M. Tourapis, K. Müller, T. Chen, “3D-TV Content Storage and Transmission,” *IEEE Trans. Broadcast*, vol. 57, no. 2, pp. 384–394, Jan, 2011. (Cité à la page 48.)
- [79] G. K. Wallace, “The JPEG still picture compression standard,” *IEEE Transactions on Consumer Electronics*, vol. 38, no. 1, pp. xviii–xxxiv, Feb. 1992. (Cité aux pages 25, 48, and 62.)
- [80] T. Acharya, P. Tsai, “JPEG2000 Standard for Image Compression”, *John Wiley & Sons*, New Jersey, 2005. (Cité à la page 48.)
- [81] S. Ouddane, SA. Fezza, K-M. Faraoun, “Stereo Image Coding : State of the Art,” in Proc. *IEEE International Workshop on Systems Signal Processing and their Applications (WoSSPA 2013)*, Tipaza, Algeria, May. 2013. pp. 122–126. (Cité aux pages 34, 48, and 68.)
- [82] M. S. Moellenhoff and M. W. Maier, “Transform coding of stereo image residuals,” *IEEE Trans. on Image Processing*, vol. 7, no. 6, pp. 804–812, Jun. 1998. (Cité à la page 48.)
- [83] W. Hachicha, A. Beghdadi, F. A Cheikh, “1D directional DCT-based stereo residual compression,” in Proc. *European signal Processing Conference (EUSIPCO 2013)*, Marrakech, Morocco, Sep. 2013. (Cité à la page 48.)
- [84] N. V. Boulgouris, M. G. Strintzis, “A family of wavelet-based stereo image coders,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 12, no. 10, pp. 898–903, 2002. (Cité à la page 48.)
- [85] J. N. Ellinas, M. S. Sangriotis, “Stereo image compression using wavelet coefficients morphology,” *Image and Vision Computing*, vol. 22, no.4, pp. 281–290, 2004. (Cité à la page 48.)

- [86] T. Frajka, K. Zeger, “Residual image coding for stereo image compression,” *Optical Engineering*, vol. 42, no. 1, pp. 182–189, Jan. 2003. (Cit     la page 48.)
- [87] A. Maalouf, M.-C. Larabi, “Bandelet-Based Stereo Image Coding,” in Proc. *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2010)*, Dallas Texas, USA, Mar. 2010, pp. 698–701. (Cit     la page 48.)
- [88] W. Woo, A. Ortega, “Overlapped block disparity compensation with adaptive windows for stereo image coding,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 10, no. 2, pp. 194–200, 2000. (Cit     la page 48.)
- [89] R. Bensalma, M.-C. Larabi, “Optimizing the disparity map by the integration of HVS binocular properties for efficient coding of stereoscopic images,” in Proc. *IEEE 3DTV-Conference (3DTV-CON 210)*, Tampere, Finland, Jun. 2010. (Cit     la page 48.)
- [90] J. N. Ellinas and M. S. Sangriotis, “Stereo Image Coder Based on the MRF Model for Disparity Compensation,” *EURASIP Journal on Applied Signal Processing*, Article ID 73950, pp. 1–13, 2006. (Cit     la page 48.)
- [91] L. B. Stelmach, W. J. Tam, D. Meegan, A. Vincent, “Stereo image quality : effects of mixed spatio-temporal resolution,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 10, no. 2, pp. 188–193, Mar. 2000. (Cit     la page 52.)
- [92] G. Saygili, C. G. Gurler, A. M. Tekalp, “Evaluation of Asymmetric Stereo Video Coding and Rate Scaling for Adaptive 3D Video Streaming,” *IEEE Trans. Broadcast.*, vol. 57, no. 2, pp. 593–601, Jun. 2011. (Cit   aux pages 51 and 55.)
- [93] Z. Liu, L. J. Karam, A. B. Watson, “Jpeg2000 encoding with perceptual distortion control,” *IEEE Transactions on Image Processing*, vol. 15, no. 7, pp. 1763–1778, 2006. (Cit     la page 59.)
- [94] Z. Wei, K. N. Ngan, “Spatial-temporal just noticeable distortion profile for grey scale image/video in DCT domain,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 19, no. 3, pp. 337–346, Mar. 2009. (Cit     la page 59.)
- [95] Y. Zhao, Z. Chen, C. Zhu, Y.-P. Tan, and L. Yu, “Binocular just-noticeable-difference model for stereoscopic images,” *IEEE Signal Process. Lett.*, vol. 18, no. 1, pp. 19–22, Jan. 2011. (Cit   aux pages 56, 57, 59, and 60.)

- [96] A. B. Watson, "DCT quantization matrices visually optimized for individual images," in Proc. *SPIE Conf. Human Vision, Visual Processing, and Digital Display*, 1993. (Cité à la page 59.)
- [97] X. Yang, W. Lin, Z. Lu, E. Ong, S. Yao, "Just noticeable distortion model and its applications in video coding," *Signal Processing : Image Communication*, vol. 20, no. 7, pp. 662-680, Aug. 2005. (Cité à la page 59.)
- [98] D. V. Meegan, L. B. Stelmach, W. J. Tam, "Unequal weighting of monocular inputs in binocular combinations : implications for the compression of stereoscopic imagery," *Journal of Experimental Psychology : Applied*, vol. 7, no. 2, pp. 143-153, Jun. 2001. (Cité aux pages 50, 51, and 52.)
- [99] L. B. Stelmach, W. J. Tam, "Stereoscopic image coding : Effect of disparate image-quality in left-and right-eye views," *Signal Process : Image Commun.*, vol. 14, no. 1, pp. 111-117, Nov. 1998. (Cité aux pages 51 and 52.)
- [100] A. Liu, W. Lin, M. Paul, C. Deng, F. Zhang, "Just Noticeable Difference for Images with Decomposition Model for Separating Edge and Textured Regions," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 20, no.11, pp. 1648-1652, Nov. 2010. (Cité à la page 58.)
- [101] R. Blake, D. H. Westendorf, R. Overton, "What is suppressed during binocular rivalry?," *Perception*, vol. 9, no 2, p. 223-231, 1980. (Cité à la page 69.)
- [102] C. Chou, C. Chen, "A perceptual optimized 3-D subband codec for video communication over wireless channels," in *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 6, no. 2, pp. 143-156, Apr. 1996. (Cité à la page 58.)
- [103] X. Zhang, W. Lin, P. Xue, "Just-noticeable difference estimation with pixels in images," *Journal of Visual Communication and Image Representation*, vol. 19, no. 1, pp. 30-41, Jan. 2008. (Cité aux pages 58 and 59.)
- [104] P. Aflaki, M. M. Hannuksela, J. Hakkinen, P. Lindroos, and M. Gabbouj, "Impact of downsampling ratio in mixed-resolution stereoscopic video," in Proc. 3DTV-Conference, Tampere, Finland, Jun. 2010. (Cité aux pages 52 and 55.)
- [105] V. De Silva, H. Arachchi, E. Ekmekcioglu, A. Fernando, S. Dogan, A. Konoz, and S. Savas, "Psycho-physical limits of interocular blur suppression and its application to asymmetric stereoscopic video delivery," in Proc. *19th International Packet Video Workshop (PV 2012)*, Munich, Germany, pp. 184-189, May. 2012. (Cité aux pages 52 and 55.)

- [106] P. Aflaki, M. M. Hannuksela, and M. Gabbouj, "Subjective quality assessment of asymmetric stereoscopic 3-D video," *Journal of Signal, Image and Video Processing*, 2013. (Cité à la page 52.)
- [107] A. Aksay, C. Bilen, E. Kurutepe, T. Ozcelebi, G. B. Akar, R. Civanlar, and A. M. Tekalp, "Temporal and spatial scaling for stereoscopic video compression," in Proc. *14th European Signal Processing Conference (EUSIPCO 2006)*, Florence, Italy, 2006. (Cité à la page 52.)
- [108] C. Fehn, P. Kauff, S. Cho, H. Kwon, N. Hur, and J. Kim, "Asymmetric coding of stereoscopic video for transmission over T-DMB," in Proc. *3DTV Conference*, Kos Island, Greece, May. 2007. (Cité à la page 52.)
- [109] Y. Chen, S. Liu, Y.-K. Wang, M. M. Hannuksela, H. Li, and M. Gabbouj, "Low-complexity asymmetric multiview video coding," in Proc. *International Conference on Multimedia and Expo (ICME 2008)*, Hannover, Germany, Apr. 2008, pp. 773-776. (Cité à la page 52.)
- [110] A. K. Jain, A. E. Robinson, and T. Q. Nguyen, "Comparing Perceived Quality and Fatigue for Two Methods of Mixed Resolution Stereoscopic Coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 24, no. 3, Mar. 2014. (Cité aux pages 52 and 54.)
- [111] A. E. Robinson, A. K. Jain, M. Scott, D. MacLeod, and T. Q. Nguyen, "Apparent sharpness of 3D video when one eye's view is more blurry," *i-Perception*, vol. 4, no. 6, pp. 456-467, 2013. (Cité aux pages 52 and 54.)
- [112] H. Brust, A. Smolic, K. Müller, G. Tech, and T. Wiegand, "Mixed Resolution Coding of Stereoscopic Video for Mobile Devices," in Proc. *3DTV-Conference*, Potsdam, Germany, 2009. (Cité à la page 52.)
- [113] H. Brust, G. Tech, K. Müller, and T. Wiegand, "Mixed resolution coding with inter view prediction for mobile 3DTV," in Proc. *3DTV-Conference*, Tampere, Finland, Jun. 2010. (Cité à la page 52.)
- [114] A. Aksay, C. Bilen, and G. B. Akar, "Subjective evaluation of effects of spectral and spatial redundancy reduction on stereo images," in Proc. *13th European Signal Processing Conference (EUSIPCO 2005)*, Antalya, Turkey, Sep. 2005. (Cité à la page 54.)
- [115] F. Shao, G. Jiang, X. Wang, M. Yu, and K. Chen, "Stereoscopic video coding with asymmetric luminance and chrominance qualities," *IEEE Transactions on Consumer Electronics*, vol. 54, no. 6, pp. 2460-2468, Nov. 2010. (Cité aux pages 54 and 55.)

- [116] S. Ince and J. Konrad, “Recovery of a missing color component in stereo images,” in Proc. *SPIE Conf. Stereoscopic Displays and Applications*, vol. 5664, San Jose, CA, USA, Jan. 2005. (Cité à la page 54.)
- [117] G. K. Wallace, “The JPEG still picture compression standard,” *IEEE Transactions on Consumer Electronics*, vol. 38, no. 1, pp. xviii–xxxiv, Feb. 1992. (Cité à la page 50.)
- [118] T. Acharya, P.-S. Tsai, “JPEG2000 Standard for Image Compression”, *John Wiley & Sons*, New Jersey, 2005. (Cité à la page 76.)
- [119] M. G. Perkins, “Data compression of stereopairs,” *IEEE Transactions on Communications*, vol. 40, no. 4, pp. 684–696, Apr. 1992. (Cité aux pages 34, 38, 41, 42, 43, and 51.)
- [120] A. Klaus, M. Sormann, K. Karner, “Segment-based stereo matching using belief propagation and a self-adapting dissimilarity measure,” in Proc. *Int. Conf. Pattern Recognition (ICPR 2006)*, pp. 15–18, Hong Kong, Aug. 2006. (Cité à la page 73.)
- [121] G. Egnal, R. P. Wildes, “Detecting binocular half occlusions : Empirical comparisons of five approaches,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 8, pp. 1127–1133, Aug. 2002. (Cité à la page 74.)
- [122] J. K. Lakshmi, M. M. Punithavalli, “A survey on skeletons in digital image processing,” in Proc. *International Conference on Digital Image Processing*, Bangkok, pp. 260–269, Mar. 2009. (Cité à la page 74.)
- [123] M. Rumpf, A. Telea, “A continuous skeletonization method based on level sets,” in Proc. *Eurographics Association of the symposium on Data Visualization*, pp. 151–155, May. 2002. (Cité à la page 74.)
- [124] D.G. Lowe, “Distinctive image features from scale-invariant keypoints,” *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, Nov. 2004. (Cité à la page 74.)
- [125] M. A. Fischler, R. C. Bolles, “Random Sample Consensus : a Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography,” *Commun. of the ACM*, vol. 24, no. 6, pp. 381–395, Jun. 1981. (Cité à la page 76.)
- [126] A. Levin, D. Lischinski, Y. Weiss, “Colorization using optimization,” *ACM Trans. Graph.*, vol. 23, no. 3, pp. 689–694, Aug. 2004. (Cité aux pages 70, 71, and 76.)

- [127] H. Hirschmuller, D. Scharstein, “Evaluation of cost functions for stereo matching,” in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2007)*, Minneapolis, MN, June 2007. (Cité aux pages 56, 62, and 76.)
- [128] G. Bjøntegaard, “Calculation of average PSNR differences between RDcurves,” in *ITU-T SG 16 Q. 6 Video Coding Experts Group (VCEG)*, Document VCEG-M33, Austin, TX, Apr. 2-4, 2001. (Cité aux pages 64, 77, and 78.)
- [129] SA. Fezza and M-C. Larabi, “Ch.10 : Color correction for stereo and multi-view coding,” in *Color Image and Video Enhancement*, Celebi, M. Emre, Lecca, Michela, Smolka, Bogdan (Eds.), pp. 291–314, Springer, 2015. (Cité aux pages 69 and 76.)
- [130] O. Dhifallah, M. Kaaniche, A. Benazza-Benyahia, “Efficient joint multiscale decomposition for color stereo image coding,” in Proc. *European signal Processing Conference (EUSIPCO 2014)*, pp. 536–540, Lisbon, Portugal, Sep. 2014. (Cité à la page 68.)
- [131] L. Cheng, S. V. N. Vishwanathan, “Learning to compress images and videos,” in Proc. *ACM 24th international conference on Machine learning (ICML)*, pp. 161–168 NY, USA, 2007. (Cité aux pages 70 and 71.)
- [132] S. Ono, T. Miyata, Y. Sakai, “Colorization-based coding by focusing on characteristics of colorization bases,” in Proc. *Picture Coding Symposium (PCS 2010)*, pp. 230–233, Nagoya, Japan, 2010. (Cité aux pages 70 and 71.)
- [133] T. Miyata, Y. Komiyama, Y. Sakai, Y. Inazumi, “Novel inverse colorization for image compression,” in Proc. *Picture Coding Symp. (PCS 2009)*, Chicago, IL, USA, May. 2009. (Cité aux pages 70, 71, and 74.)
- [134] T. Ueno, T. Yoshida, M. Ikehara, “Color image coding based on the colorization,” in Proc. *Asia-Pacific Signal & Information Processing Association Annual Summit and Conference (APSIPA ASC)*, CA, USA, December 2012. (Cité à la page 70.)
- [135] T. Horiuchi, S. Tominaga, “Color image coding by colorization approach,” *EURASIP Journal on Image and Video Processing*, vol. 1, pp. 1–9, 2008. (Cité aux pages 70 and 71.)
- [136] S. Lee, S.-W. Park, P. Oh, M.G. Kang, “Colorization-based compression using optimization,” *IEEE Trans. on Image Processing*, vol. 22, no. 7, pp. 2627–2636, Jul. 2013. (Cité aux pages 70, 71, and 74.)

- [137] K. Mishiba, T. Yoshitome, “Colorization matrix construction with high compression efficiency for colorization-based coding using optimization,” in *Proc. IEEE International Conference on Image Processing (ICIP2014)*, pp.5551–5555, Paris, France, October 2014. (Cit  aux pages 70 and 71.)
- [138] P. Oh, S. Lee, M. G. Kang, “Local regression based colorization coding,” in *Proc. International Conference on Computer Vision Theory and Applications (VISAPP)*, pp.153–159, Lisbon, Portugal, Jan. 2014. (Cit  aux pages 70 and 71.)
- [139] P. Oh, S. Lee, M. G. Kang, “Colorization coding based on iterative regression with major group priority,” *Digital Signal Processing*, vol. 56, pp. 53–66, 2016. (Cit  aux pages 70 and 71.)
- [140] K. Uruma, K. Konishi, T. Takahashi, T. Furukawa, “Fast colorization based image coding algorithm using multiple resolution images,” *EURASIP Journal on Image and Video Processing*, vol. 1, pp. 1–15, 2016. (Cit  aux pages 70 and 72.)
- [141] B. Julesz, *Foundations of Cyclopean Perception*. Univ. Chicago Press, 1971. (Cit    la page 69.)
- [142] A. Popowicz, B. Smolka, ”Overview of Grayscale Image Colorization Techniques”, in *Color Image and Video Enhancement, Springer International Publishing*, pp. 345–370, 2015. (Cit    la page 70.)
- [143] S. Ouddane and K-M. Faraoun, “Asymmetric Stereoscopic Images Coding using Perceptual Model,” *Journal of Multimedia Processing and Technologies*, vol. 9, no. 2, pp. 59–70, 2018. (Cit  aux pages 49 and 55.)
- [144] S. Ouddane, K-M. Faraoun, SA. Fezza and M-C. Larabi, “Adaptive colorization-based compression for stereoscopic images,” in *Proc. IEEE 3DTV-Conference : The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON’2016)*, Hamburg, Germany, 2016. (Cit    la page 69.)