



HAL
open science

Estimer, mesurer et corriger les artefacts de compression pour la télévision numérique

Frédérique Crété-Roffet

► **To cite this version:**

Frédérique Crété-Roffet. Estimer, mesurer et corriger les artefacts de compression pour la télévision numérique. Traitement du signal et de l'image [eess.SP]. Université Joseph-Fourier - Grenoble I, 2007. Français. NNT: . tel-00229416

HAL Id: tel-00229416

<https://theses.hal.science/tel-00229416>

Submitted on 31 Jan 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université Joseph Fourier

No. attribué par la bibliothèque

--	--	--	--	--	--	--	--	--	--	--	--

THESE

pour obtenir le grade de

DOCTEUR DE L'UJF

Spécialité : Signal - Image - Parole - Télécom

préparée au Gipsa-lab, laboratoire Grenoble, image, parole, signal et automatique.

dans le cadre de l'Ecole Doctorale :

Electronique, Electrotechnique, Automatique, Traitement du Signal

présentée et soutenue publiquement

par

Frédérique CRETE-ROFFET

le 29 novembre 2007

Titre :

**ESTIMER, MESURER ET CORRIGER LES ARTEFACTS DE COMPRESSION
POUR LA TÉLÉVISION NUMÉRIQUE**

JURY

Michel Jourlin	Président du jury	Professeur/ISTASE/Saint-Etienne
Christine Fernandez-Maloigne	Rapporteur	Professeur/SIC/Poitiers
Touradj Ebrahimi	Rapporteur	Professeur/EPFL/Lausanne
Patrick Le Callet	Examineur	Maître de conférences/IRCCyN/Nantes
Jeanny Hérault	Directeur de thèse	Professeur/GIPSA-Lab/Grenoble
Patricia Ladret	Co-Directrice de thèse	Maître de conférences/GIPSA-Lab/Grenoble
Marina Nicolas	Encadrante entreprise	Ingénieur R&D/STMicroelectronics/Grenoble

Remerciements

Même si ce chapitre est, de loin, le moins stressant à écrire, je lui accorde autant d'importance que les autres car il me permet de remercier toutes les personnes qui ont participé à cette aventure.

Je commencerai par adresser mes remerciements aux personnes du jury qui m'ont fait l'honneur d'accepter d'évaluer cette thèse.

Monsieur Michel Jourlin, pour votre enseignement du traitement d'images dans les années 2002 et 2003 qui m'a donné l'envie de poursuivre dans ce domaine.

Madame Christine Fernandez-Maloigne et Monsieur Touradj Ebrahimi, pour vos remarques pertinentes et constructives dans le domaine de la qualité et de la compression qui ont contribué à l'amélioration du contenu de cette thèse et ont donné des éléments de réponses aux perspectives évoquées.

Monsieur Patrick Le Callet, pour avoir relevé certains points souvent mis de côté, et pour vos conseils pendant la thèse et pour la soutenance.

Je tiens évidemment à exprimer mes remerciements aux deux personnes qui m'ont suivie pendant ces trois années, Patricia et Marina. Merci pour votre encadrement, votre confiance, votre sympathie et votre soutien que j'ai pu apprécier et trouver à la fois dans le domaine universitaire et dans le domaine industriel. Cet équilibre m'a permis de ne pas négliger un domaine par rapport à l'autre. Merci également à Jeanny qui nous a accordé toute sa confiance pour ce travail de thèse.

J'adresse aussi mes remerciements au Professeur Roger-David Hersch de l'EPFL, qui m'a accueillie dans son laboratoire lors de mon année de DEA et qui m'a énormément appris sur la Recherche.

Merci également aux équipes IQI de STMicroelectronics et GPiG du Gipsa-lab dans lesquelles j'ai pu évoluer et enrichir mes connaissances tout au long de ces 3 années de thèse.

Jérôme et Pierre, je ne pourrai pas oublier ces moments partagés dans le box visiteur. Le Truman Show est fini mais j'aurais bien voulu un exemplaire pour re-visionner certains passages. Merci pour ces délires, je pense qu'on a du faire exploser l'audience plusieurs fois!

Je désire adresser maintenant un merci plus particulier à Grégory. Tu as su me soutenir pendant ces trois années et surtout ces derniers mois d'écriture..., de relecture..., et d'entraînement pour la soutenance. Tu as toujours trouvé les mots pour me re-motiver, MERCI pour tout !

Naturellement, merci à vous 2, Maman et Michel. Vous avez toujours su me mettre dans les meilleures conditions pour que je poursuive mes études. Je crois que vous ne vous attendiez pas à ce que cela dure aussi longtemps!!! Olivier et Benoit, encore merci d'avoir pris vos places Dern's pour me faire cette agréable surprise le jour de la soutenance. Et je n'oublie pas tous les membres de ma famille qui se reconnaîtront, qui m'ont toujours soutenue et qui ont toujours porté un intérêt à ce que je faisais.

Enfin merci à celles et ceux qui sont venus assister (ou qui auraient souhaité être là si il n'y avait pas eu toute cette distance) à mes pré-soutenances, soutenance, pot et soirée de thèse !

C'était une très belle aventure, inoubliable, mais voilà, toutes les bonnes choses ont une fin !
Ouf...

A Dominique.

Table des matières

Liste des figures	11
Liste des tableaux	17
Introduction générale	19
1 Contexte	19
2 Définition du sujet	20
3 Organisation du document	21
Partie I Analyse du sujet et Problématique	23
Chapitre 1 La Compression Numérique : Rôle, Principe et Conséquences	25
1 Les normes de compression	26
1.1 Les organismes de normalisation	26
1.2 Les normes spécifiques aux images fixes	26
1.3 De l'image à la vidéo	27
1.4 Les normes spécifiques à la vidéo	27
1.5 Un bilan sur les normes et les applications	29
1.6 Discussion	30
2 Les principes de la compression vidéo	31
2.1 Le flux vidéo	31
2.2 La compression spatiale	32
2.3 La compression temporelle	37
3 Les artefacts de compression	41
3.1 Les distorsions spatiales	41
3.2 Les distorsions temporelles	47
Chapitre 2 Etat de l'art des corrections	51
1 Le choix des artefacts et de la catégorie de correction	52
1.1 Le choix des artefacts	52
1.2 Le choix de la catégorie de correction	53
1.3 Les méthodes de pré-traitements avant l'encodage	53
1.4 Les méthodes liées à l'encodage	53

1.5	Les méthodes post-traitements après le décodage de l'image	54
2	Etat de l'art des méthodes d'amélioration d'images	57
2.1	Des méthodes inhabituelles	57
2.2	Des techniques de filtrage linéaires et non linéaires	58
2.3	Les corrections par classification et filtrage	61
3	Les limitations des méthodes de classification et filtrage	66
3.1	L'apparition des fausses frontières	66
3.2	La confusion entre frontières de blocs et contours de l'image	69
3.3	L'apparition du flou	71
 Bilan de la première partie		 75
 Partie II La mesure des artefacts		 77
 Chapitre 3 L'estimation de la qualité		 79
1	Bien définir le terme de qualité	80
1.1	La notion de fidélité	80
1.2	La notion de qualité	80
2	Les mesures subjectives	81
2.1	Un protocole bien défini	81
2.2	Les tests psychophysiques sur les images fixes	82
2.3	Les tests psychophysiques sur les séquences animées	84
2.4	La note objective de qualité	85
2.5	Les contraintes des tests subjectifs	86
3	Les mesures objectives de fidélité	88
3.1	Les mesures de distances	88
3.2	Les critères issus du traitement du signal	89
3.3	La problématique de ces mesures de fidélité	89
4	Les mesures de qualité objectives	91
4.1	Les différentes catégories de métriques de qualité objectives	91
4.2	La validation des métriques	96
4.3	Une mesure de qualité globale ou spécifique	97
5	Retour au contexte	99
5.1	Choix de la catégorie de métrique	99
5.2	La stratégie envisagée	99
 Chapitre 4 La mesure de l'effet de bloc		 103
1	Etat de l'art des mesures de bloc sans référence	104
1.1	Dans le domaine fréquentiel	104
1.2	Dans le domaine spatial	106
2	Le BLE : Block Level Estimator	113
2.1	La perception de l'effet de bloc	113
2.2	Attribuer un coefficient traduisant la gêne de l'effet de bloc	116
2.3	Différencier l'effet de bloc du contenu de l'image	118

2.4	Cas des forts taux de compression	120
2.5	L'échelle du BLE	123
3	La validation du BLE	126
3.1	Les tests subjectifs	126
3.2	La corrélation entre le MOS et les tests subjectifs	127
3.3	La confrontation du BLE avec les autres métriques	128
4	Les limitations du BLE	132
4.1	La position des frontières	132
4.2	L'aspect temporel	133
4.3	Le masquage de luminance et de la chrominance	133
Chapitre 5 La mesure du flou		135
1	Etat de l'art des différentes mesures de flou et de leurs applications	136
1.1	Le flou causé par une mauvaise focalisation	136
1.2	Le flou gênant pour le SVH	137
2	Proposition d'une métrique de flou : le BluM ou BlurMetric	140
2.1	La discrimination du flou	140
2.2	Le principe de la métrique	141
2.3	Les caractéristiques à prendre en compte	143
2.4	L'implémentation de la métrique	144
3	Validation du BluM et résultats	147
3.1	Les tests subjectifs	147
3.2	L'analyse des tests subjectifs	148
3.3	La comparaison du BluM avec la métrique de <i>Marziliano et al</i>	154
3.4	La robustesse au bruit	155
4	Le rôle de la couleur dans la perception du flou	158
4.1	L'étude subjective	158
4.2	L'étude objective	159
Bilan de la seconde partie		165
Partie III La correction des artefacts		167
Chapitre 6 Proposition de modules d'amélioration de la qualité des images compressées		169
1	Le BLE pour contrôler la correction et éviter toute confusion	170
1.1	Problématique et domaine d'application	170
1.2	La solution proposée	170
1.3	L'efficacité de la méthode	174
2	La suppression des frontières fantômes	176
2.1	Rappel de la problématique	176
2.2	Domaine d'application	176
2.3	Les solutions proposées	176

2.4	Les méthodes idéales de détection et de correction des frontières de blocs entre zones homogènes	177
2.5	L'intégration dans un algorithme temps réel	179
2.6	Diagramme Bloc de l'algorithme	183
2.7	Les résultats	183
3	La correction proche des contours et dans les zones texturées	187
3.1	Domaine de définition	187
3.2	Rappel de la problématique	187
3.3	La solution proposée	189
3.4	La prise en compte de l'effet de bloc	191
3.5	Les résultats	192

Chapitre 7 Proposition et évaluation d'un algorithme de suppression des artefacts de compression MPEG-2 **195**

1	Description de l'algorithme	196
1.1	Domaine d'application	196
1.2	Description du diagramme bloc	196
2	L'évaluation des performances	203
2.1	L'efficacité du comportement du SQDSE	203
2.2	L'efficacité globale du SQDSE	210
3	Les extensions	215
3.1	L'intégration d'une correction des fausses couleurs	215
3.2	L'automatisation de la recherche de taille des MBs et du décalage	215
3.3	Les tests subjectifs sur des séquences	216

Bilan de la troisième partie **219**

Conclusion générale **221**

Glossaire **225**

Publications et Brevets **227**

Bibliographie **229**

Résumé **237**

Liste des figures

1.1	Hiérarchie dans le flux vidéo.	31
1.2	Sélection des pixels de chrominance dans le cas d'un format 4 :2 :0.	32
1.3	Découpage d'une image en Macroblocs et en Blocs.	33
1.4	Catalogue des motifs d'un bloc DCT 8x8.	34
1.5	Exemples de passage des valeurs des pixels dans le domaine fréquentiel avec la DCT pour un bloc homogène et un bloc non-homogène.	35
1.6	Effet de la quantification sur les coefficients DCT.	35
1.7	Balayage en zig zag de la matrice DCT quantifiée.	36
1.8	Exemple d'organisation d'un GOP de 12 images.	37
1.9	De gauche à droite : Informations à coder pour une image I, une image P et une image B.	38
1.10	Principe de l'estimation de mouvement.	39
1.11	L'effet de bloc.	41
1.12	Effet de masquage : l'effet de bloc est plus visible dans les zones avec peu de détails que dans les zones contenant des variations d'informations.	42
1.13	L'effet de motif.	42
1.14	Le flou.	43
1.15	La décomposition d'une image avec les composantes Y, Cb et Cr.	44
1.16	Les fausses couleurs.	45
1.17	L'effet d'escalier.	45
1.18	L'effet de ringing.	46
1.19	Les faux contours.	47
1.20	L'effet de moustique.	48
1.21	Exemple de disparités dues à la compensation de mouvement.	49
2.1	Principe de la méthode de Kutka et al : pré-traitement par rehaussement du contraste	53
2.2	Principe général des méthodes de restauration.	55
2.3	Cheminement suivi pour choisir la catégorie de méthode de correction.	56
2.4	Schématisation des pixels d'une image I et des coordonnées de ces pixels dans un voisinage 3*3	58
2.5	Exemple de correction de type filtrage et rehaussement. α est le coefficient du filtre de Jarske et al dont le noyau est représenté à l'équation (2.6).	60
2.6	Les limites du filtrage dans une région définies par Meier et al : le filtrage du pixel considéré ne sera effectué que dans le cas a).	61

2.7	a) Représentation de différents contenus dans une paire de blocs. b) Représentation des zones de filtrage qui s'adaptent au contenu des blocs.	62
2.8	Les filtres de Robinson	63
2.9	Exemple de filtrage passe-bas au niveau des frontières de blocs : méthodes de Lin et al	64
2.10	Cheminement suivi pour choisir la méthode de correction pour la réduction de l'effet de bloc et de l'effet de ringing.	65
2.11	Principe traditionnel appliqué pour la correction de l'effet de bloc.	66
2.12	Le phénomène de création de fausses frontières ou frontières fantômes notées G_{n_G} et G_{n_D}	67
2.13	Le phénomène des blocs étendus : les frontières de blocs sont moins nombreuses sur l'image de droite qui est pourtant plus compressée (0.4 bits par pixels) que l'image de gauche (0.6 bits par pixels)	67
2.14	De gauche à droite : image non compressée, image compressée , image corrigée par filtre moyennneur sur 4 pixels à gauche et à droite de la frontière (les frontières fantômes sont visibles).	68
2.15	La confusion entre les frontières de blocs et les contours de l'image.	69
2.16	De gauche à droite : la zone de l'image est de plus en plus compressée et montre l'évolution de la dégradation du contour.	70
2.17	Problème du seuillage avec les filtres passe-haut. A : Image originale, B : Contours verticaux détectés par un filtre de Sobel. C : les intensités au dessus de 100 sont considérées comme des contours D : idem avec une valeur de seuil égale à 200.	72
2.18	Les artefacts de correction : conséquences du traitement des artefacts de correction.	73
2.19	Bilan de la première partie.	76
3.1	Les tests de Snellen et d'Ichihara.	81
3.2	Le test d'ordonnancement.	83
3.3	Le protocole des méthodes à simple stimulus.	84
3.4	A gauche : échelle de notation continue. A droite : échelle de notation par catégories de qualité.	84
3.5	Le protocole des méthodes à double stimuli.	84
3.6	L'échelle comparative.	85
3.7	Défaut du PSNR et du MSE dans l'estimation de la qualité. Par rapport à l'image originale, les deux images bruitées ont une qualité perçue différente mais un PSNR et un MSE très proche.	90
3.8	La sensibilité au contraste en fonction de la fréquence spatiale selon [Van Nes et Bouman, 1967], cpd : cycles par degré.	92
3.9	Exemple d'approche d'un processus ascendant : l'image est modifiée pour mettre en avant ce que le SVH perçoit.	96
3.10	Figure extraite tirée de [Wang et al., 2004], qui représente la corrélation entre le MOS et le SSIM pour deux bases de données d'images.	97
3.11	A gauche : Zoom sur une image compressée, les frontières de bloc sont visibles. A droite, zoom sur une image corrigée par filtrage : les frontières de bloc ne sont plus visibles mais l'image est floue.	100
4.1	Principe de mesure d'effet de blocs par Wang et al [Wang et al., 2000].	104

4.2	Images-tests que nous utilisons pour l'évaluation de la métrique de Wang et al [Wang et al., 2000].	105
4.3	Résultats de la métrique de Wang et al [Wang et al., 2000] et du PSNR sur les images présentées sur la figure 4.2.	105
4.4	Représentation d'un bloc 8*8 et de ses blocs voisins. (Figure extraite de [Pan et al., 2004])	107
4.5	Possible confusion entre une frontière et un contour : 1) Représentation de deux paires de blocs de contenus différents. 2) Mise en évidence des variations d'informations pour chaque ligne de ces blocs. 3) Somme des résultats obtenus sur chaque ligne.	108
4.6	Les éléments en gris définissent les ensembles Ω_{1V} , Ω_{2V} et Ω_2 dans un bloc de 8*8 pixels qui a été convolué à un filtre de Sobel.	109
4.7	La frontière entre la paire de blocs de gauche ets plus visible que la frontière entre la paire de blocs de droite. Pourtant, ces deux frontières sont vues de la même façon selon la normalisation de Perra et al.	109
4.8	Diagramme bloc de la métrique de Tan et al.	110
4.9	Exemple de mise en valeur des frontières de blocs verticales par l'application d'un filtre de Sobel et par soustraction des contours de l'image décodée aux contours de l'image originale.	111
4.10	Les raisons de la perception d'une frontière de bloc.	113
4.11	Les étapes de la modélisation déterminant la visibilité d'une frontière.	114
4.12	La représentation binaire des frontières de bloc : visible ou non-visible.	115
4.13	L'intérêt de prendre en compte le phénomène de masquage : l'information uniquement au niveau de la frontière n'est pas corrélée à notre perception.	116
4.14	Les étapes de la modélisation déterminant le coefficient représentatif de la gêne d'une frontière.	117
4.15	La représentation des différents degrés de visibilité d'une frontière.	118
4.16	A gauche : composante d'intensité de l'image "vitrail". A droite : composante d'intensité de l'image "vitrail" fortement comprimée. L'écart maximum trouvé entre les deux niveaux de gris de la frontière est de 32.	119
4.17	La compression a des répercussions visuelles que sur les zones contenant des variations d'informations.	120
4.18	Le phénomène des blocs étendus.	120
4.19	Image type où la confusion est possible entre les zones homogènes de l'image et des blocs étendus.	121
4.20	Définition et schématisation d'un bloc étendu.	122
4.21	A) Image I1 moyennement comprimée, Image I2 fortement comprimée ; B) Représentation de l'état binaire des frontières sans tenir compte des blocs étendus ; C) Représentation de l'état binaire des frontières en tenant compte des blocs étendus ; D) Représentation du degré de visibilité des frontières en tenant compte des blocs étendus ;	124
4.22	Echantillon des images présentées dans le test subjectif.	126
4.23	Corrélation entre le MOS et le BLE.	127
4.24	Niveaux de dégradation associés aux valeurs de la métrique.	127
4.25	La métrique de Wang et al M1 vs le MOS.	128

4.26	La métrique de Perra et al M2 vs le MOS.	128
4.27	Le BLE indique que l'effet de bloc est gênant sur l'image compressée de gauche et que l'effet de bloc est imperceptible sur l'image floue de droite.	134
5.1	La détection du flou sur <i>Léna</i> par les méthodes de [Elder et Zucker, 1998] et [Hu et Haan, 2006] : les pixels clairs indiquent une zone floue	137
5.2	Diagramme bloc de la méthode de Marziliano et al [Marziliano et al., 2002].	138
5.3	Principe de détection des points d'inflexion. Image tirée de [Marziliano et al., 2004] 138	
5.4	Expérience sur la discrimination du flou : 1) images à classer de la plus nette à la plus floue. 2) Résultat.	140
5.5	Histogrammes des variations entre pixels voisins sur une image nette, floutée et re-floutée.	141
5.6	Principe de la métrique.	142
5.7	Illustration du principe de la métrique sur une image nette.	142
5.8	Illustration du principe de la métrique sur une image floue.	142
5.9	La perception du flou suivant le contenu de l'image.	143
5.10	Les pertes et les gains des intensités des pixels après avoir flouté d'une image.	143
5.11	La perception du flou unidirectionnel.	144
5.12	Description des étapes de l'algorithme de la mesure du flou.	144
5.13	Echelle de perception du BluM.	146
5.14	La salle des tests subjectifs.	147
5.15	Echantillon des images présentées aux observateurs.	148
5.16	Temps de réponse moyen pour l'échelle des notes.	149
5.17	Corrélation entre le MOS et le BluM.	150
5.18	Approximation linéaire pour déterminer les coefficients G et D_M	151
5.19	Attribution d'un degré de visibilité du flou en fonction des valeurs de la métrique.	152
5.20	Résultats de la métrique de flou sur différentes images.	153
5.21	Robustesse de la métrique selon le contenu.	154
5.22	Corrélation de la métrique avec notre limite de perception du flou.	154
5.23	Corrélation de la métrique de <i>Marziliano et al</i> avec le MOS.	155
5.24	Corrélation entre les notes subjectives et le BluM et la métrique de <i>Marziliano et al</i> en fonction du flou utilisé.	155
5.25	L'analyse de la robustesse de la métrique au bruit analogique et au bruit numérique.	156
5.26	La perception du flou sur des images avec et sans couleur.	158
5.27	Une image couleur décomposée en Y, Cb et Cr	159
5.28	A droite : diagramme de chromaticité avec position de quelques couleurs. A gauche : Ellipses de MacAdam dans le diagramme de chromaticité	160
5.29	L'adaptation de la métrique à la couleur.	161
5.30	Corrélation entre le BluM adapté à la couleur et le MOS.	161
5.31	La métrique de bloc et la métrique de flou pour évaluer la qualité d'une correction.	163
5.32	Bilan de la seconde partie.	166
6.1	Rappel sur les causes des artefacts de correction.	169
6.2	La distinction des IPB par le BLE sur la séquence <i>swim</i>	171
6.3	Résultat du filtrage F1 pour des blocs compressés à taux moyen.	172
6.4	Résultats des filtrages F2 et F1 pour des blocs fortement compressés.	173

6.5	Diagramme bloc du module d'amélioration intégrant le BLE pour le contrôle de la correction.	173
6.6	Les conséquences de la confusion entre les frontières de blocs et les contours et l'intérêt d'intégrer le contrôle du BLE.	174
6.7	L'intérêt d'utiliser le BLE pour contrôler la force de filtrage	175
6.8	A gauche : la problématique liée à l'intégration temps réel. A droite : les répercussions avec l'apparition des frontières fantômes	176
6.9	La solution idéale pour supprimer les frontières fantômes.	177
6.10	A gauche : les blocs étendus après une forte compression. A droite : Proposition de correction par extension de la zone de filtrage.	177
6.11	Résultat obtenu par étalement du dégradé sur les deux blocs et comparaison avec les méthodes classiques.	178
6.12	Résultat obtenu par étalement du dégradé sur le bloc étendu et comparaison avec les méthodes classiques.	178
6.13	La première étape de correction avant l'étape de détection.	179
6.14	La correction de la frontière suivante non biaisée par la précédente.	180
6.15	La correction de la frontière suivante non biaisée par la précédente dans le cas d'un bloc étendu.	180
6.16	La comparaison des différentes méthodes de correction en présence d'un bloc étendu.	181
6.17	Correction dans le cas où les niveaux de gris des blocs ne suivent pas un ordre.	182
6.18	Diagramme bloc d'une méthode de correction commentant notre module d'amélioration pour éviter le problème des fausses frontières.	183
6.19	Image I1 originale, compressée, corrigée avec des méthodes traditionnelles, corrigée avec notre méthode	184
6.20	Image I2 originale, compressée, corrigée avec des méthodes traditionnelles, corrigée avec notre méthode	184
6.21	Rappel des échelles des différentes métriques.	185
6.22	Schématisation d'un contour de faible intensité et d'un contour de forte intensité entourés de bruit.	187
6.23	Contour de faible intensité assimilable à de l'effet de ringing proche d'un contour de forte intensité.	188
6.24	Effet de bloc entre un bloc contenant un contour et un bloc homogène et effet de ringing le long du contour.	188
6.25	Effet de bloc entre deux blocs contenant un contour et effet de ringing le long du contour.	189
6.26	Effet de bloc entre deux blocs texturés.	189
6.27	Diagramme bloc du principe de correction de l'effet de ringing.	191
6.28	Diagramme bloc du principe de correction de l'effet de ringing et de l'effet de bloc.	192
6.29	Zoom d'un résultat de la correction du FIV sur des contours de différentes intensités. Image de gauche compressée, Image de droite corrigée	192
6.30	Modules de suppression de l'effet de bloc et de l'effet de ringing.	194
7.1	Image compressée et image compressée redimensionnée et décalée.	196
7.2	Diagramme Bloc de l'algorithme SQDSE.	197

7.3	Cas d'une frontière entre deux blocs dans le mode progressif et dans le mode entrelacé.	198
7.4	Cas d'un bloc réel encodé en entrelacé (à gauche) et en progressif(à droite).	198
7.5	Détection du mode progressif et du mode entrelacé.	199
7.6	Les frontières traitées et leur ordre de traitement.	199
7.7	Cas d'une frontière horizontale entre deux MBs codés en mode entrelacé.	200
7.8	Cas des frontières horizontales : transposition de la matrice de traitement pour traiter comme une frontière verticale.	200
7.9	Conséquences de l'estimation du BLE sur l'image précédente.	203
7.10	Conséquences de l'estimation du BLE sur l'image précédente sur le flou.	204
7.11	BLE sur la séquence swim, la séquence corrigée par le <i>MSU Video Group</i> et la séquence corrigée par le SQDSE.	204
7.12	BLE sur la séquence canada, la séquence corrigée par le <i>MSU Video Group</i> et la séquence corrigée par le SQDSE.	205
7.13	Effet des corrections <i>MSU Video Group</i> et SQDSE sur des MBs codés en entrelacé. 206	
7.14	Mesure du flou sur les séquences compressées et corrigées par le <i>MSU Video Group</i> et le SQDSE.	207
7.15	L'efficacité du SQDSE après un décalage de l'image.	208
7.16	L'efficacité du SQDSE après un décalage et un changement de dimension de l'image. 208	
7.17	Impact du SQDSE sur l'effet de moustique : à gauche : image différence entre 2 images consécutives sans correction; à droite : image différence entre 2 images consécutives avec correction.	209
7.18	BLE et BluM moyen sur 40 séquences de différents contenus et plus ou moins compressées en MPEG-2 et corrigées avec le SQDSE.	210
7.19	Image de test pour l'évaluation du SQDSE.	211
7.20	Efficacité du SQDSE face au respect de contenu et à la suppression de frontières prononcées.	211
7.21	Efficacité du SQDSE face au respect de la texture.	212
7.22	Efficacité du SQDSE face au respect des contours.	212
7.23	Image <i>Lilo</i> compressée et corrigée par le SQDSE.	213
7.24	Image <i>Canada</i> compressée et corrigée par le SQDSE.	214
7.25	L'impact de la correction de l'effet de bloc et de l'effet de ringing sur la perception des fausses couleurs.	215
7.26	L'efficacité du SQDSE.	217
7.27	Bilan de la seconde partie.	220

Liste des tableaux

1.1	Les normes et leurs applications	30
1.2	Sous-échantillonnage des composantes Cb et Cr	32
1.3	Liste des artefacts causés par la compression spatiale et/ou temporelle, apparaissant au niveau du bloc et/ou du macrobloc	49
2.1	Les différents filtrages proposés par Lee et al	70
3.1	Echelle de notation.	83
4.1	BLE, M1 et M2 sur les images de la figure 4.17	129
4.2	BLE, M1 et M2 sur les images de la figure 4.18	129
6.1	Récapitulatif des mesures obtenues par différentes métriques sur l'images I1 : A : non comprimées, B : comprimées, C : corrigées avec des correction classiques, D : corrigées avec notre correction	185
6.2	Récapitulatif des mesures obtenues par différentes métriques sur l'image I2 : A : non comprimées, B : comprimées, C : corrigées avec des correction classiques, D : corrigées avec notre correction	185

Introduction générale

1 Contexte

L'apparition de l'informatique a favorisé le passage de l'analogique au numérique. Les premières données ont été numérisées par l'intermédiaire d'un code binaire (le bit généralement noté 0 ou 1). Cette solution a vite été adoptée car, contrairement à un signal électrique, elle peut être reproduite indéfiniment, stockée plus rigoureusement et modifiée beaucoup plus facilement. De nombreuses applications ont vu le jour comme les appareils photographiques numériques, les caméscopes numériques et plus particulièrement la télévision numérique.

La télévision numérique La télévision numérique s'est largement développée par l'intermédiaire de trois types de réseaux de diffusion :

- la distribution directe par satellite (par exemple, Canal Satellite ou TPS).
- la distribution par les émetteurs hertziens (la TNT : Télévision Numérique Terrestre).
- la distribution par câble (par exemple la ligne ADSL).

Ce signal peut être vu sur un téléviseur analogique, numérique, un écran d'ordinateur ou encore sur un téléphone portable à condition d'avoir les décodeurs appropriés.

La transmission du signal Qu'il soit transmis dans les airs ou par câble, le signal doit être codé. Ce code contient les traitements appliqués au signal qui sera transmis. Ce n'est qu'en connaissant ce descriptif que le décodeur, situé à la réception, aura les moyens de retrouver et de reconstruire le signal transmis. Cela explique la nature des différents décodeurs en fonction de l'émetteur.

Quelques ordres de grandeur Pour diffuser en numérique du signal TV en définition standard et de qualité équivalente à l'analogique, il faudrait délivrer continûment dans le réseau de diffusion quelques $576 \times 720 \times 25 = 10 \text{ Mbits/s}$, 576 étant le nombre de lignes par image, 720 étant le nombre de pixels contenu dans chaque ligne et 25 étant le rythme de défilement d'images par seconde. Ce calcul empirique suppose en plus que chaque pixel soit codé sur 1 bit ce qui n'est évidemment pas le cas, car chaque pixel d'image contient des informations de lumière et de couleur qui nécessitent plus d'un bit pour être codé. On est donc dans des débits de l'ordre de 150 Mbits/s pour des images TV de définition standard.

Or, la bande passante attribuée pour la TNT correspond à un débit maximal de 24 Mbits/s, 36 Mbits/s pour le satellite. Des débits plus importants peuvent être atteints avec les lignes téléphoniques ADSL mais la qualité dépendra alors de la distance qui sépare la ligne du central de réception le plus proche.

La compression On comprend alors pourquoi la compression des données est indispensable à la diffusion des chaînes numériques. La compression peut être de deux types :

- La compression dite sans pertes effectuant un traitement totalement transparent, permettant de retrouver intégralement les données d'origine à l'issue de la décompression mais offrant des taux de compression faibles.
- La compression avec pertes délivrant des taux de compression nettement supérieurs mais imposant de négliger certaines informations de l'image en tenant compte de sa nature et de notre perception visuelle. Si elle se fait dans des proportions limitées, l'élimination de ces informations peut passer inaperçue pour un observateur moyen. En revanche, si la compression est réalisée dans des proportions élevées, la perte de données se traduira par l'apparition de distorsions plus ou moins visibles.

Normaliser le codage et la compression Avant d'être transmis, le signal doit donc non seulement être codé mais aussi compressé. A la réception, le signal sera restitué par le décodeur approprié. Pour favoriser le transport universel des données, ces principes sont décrits dans des normes. La norme de codage majoritairement utilisée en télévision numérique est le MPEG-2. Pour donner un ordre de grandeur, une chaîne TV de définition standard peut être codée en utilisant 3 à 6 Mbits/s permettant ainsi l'envoi de plusieurs chaînes.

2 Définition du sujet

Enoncé du sujet Les débits dédiés à la diffusion numérique ne sont atteints qu'en utilisant une compression avec pertes. Ces pertes peuvent se traduire par l'apparition de dégradations sur les images diffusées. Ces dégradations, plus communément appelées artefacts vont déterminer l'impact de la compression sur l'image perçue sur les téléviseurs.

Le travail demandé : "Réduction des artefacts de compression MPEG-2" consiste à trouver des solutions permettant d'améliorer la qualité des images diffusées. Cette solution devra être intégrée aux décodeurs ou directement dans les téléviseurs.

Les objectifs Pour réaliser ce travail, deux objectifs sont ciblés :

- D'une part, comment quantifier l'impact des artefacts de compression sur la qualité d'une image perçue ?
- D'autre part, peut-on minimiser cet impact en évitant, contrôlant ou supprimant ces artefacts ?

Pour atteindre ces objectifs, le travail de recherche sera effectué en collaboration avec le gipsa-lab Département Images Signaux et l'équipe de recherche et développement Image Quality Improvement de STMicroelectronics et devra répondre à des contraintes pour aboutir à un produit exploitable par l'entreprise STMicroelectronics qui développe et produit des circuits à base de semi-conducteurs.

Les contraintes Nous distinguons quatre catégories de contraintes.

Premièrement, la solution est destinée à être intégrée sur un circuit pour des applications grand public. Sa commercialisation à grande échelle dépendra alors de son prix, c'est-à-dire principalement de la taille de silicium nécessaire à sa réalisation. Or, plus les calculs sont complexes

et nombreux, plus la taille de silicium augmente. Nous devons donc aboutir à un principe intégrant une architecture la plus simple possible pour répondre à cette contrainte de "faible coût".

Cette puce sera intégrée dans les décodeurs ou encore directement dans les téléviseurs et devra améliorer la qualité des images reçues avant de les afficher. Pour être efficace, la solution devra alors fonctionner rapidement pour éviter un décalage de l'affichage des images. Cette contrainte est plus communément appelée "la contrainte de temps réel".

Pour que la solution soit choisie par le grand public, elle doit être efficace sur la majorité des images diffusées et ne doit jamais dégrader. En effet, il est relativement simple d'améliorer la qualité d'une image précise mais il est beaucoup plus difficile de rendre cette solution efficace dans tous les cas. La troisième contrainte est donc de trouver un compromis aboutissant à une solution efficace dans la majorité des cas et non dégradante dans les autres cas, nous l'appellerons la contrainte de "robustesse".

Enfin, en vue d'élargir son domaine d'application, la solution doit être la plus indépendante possible et ne doit pas utiliser les informations données par le décodeur renseignant par exemple sur la force de compression. Nous partirons avec l'hypothèse que nous traitons un signal encodé en MPEG-2 mais, si ce n'est pas le cas, notre solution devra s'adapter et éviter d'introduire de nouvelles dégradations.

3 Organisation du document

Ce document est décomposé en trois grandes parties dont nous allons décrire le contenu.

Le rôle de la première partie est de faire l'état de l'art sur le sujet et de définir la problématique.

Première partie

Chapitre 1 A partir d'un historique sur les différentes normes, nous analyserons précisément la norme MPEG-2 utilisée pour la diffusion numérique. Cette analyse permettra de comprendre la formation des différents artefacts et leur impact sur notre perception.

Chapitre 2 Nous étudierons alors les méthodes proposées dans la littérature pour corriger ces artefacts. En ciblant les corrections qui répondent à nos contraintes, nous verrons que la correction des artefacts de compression se fait parfois au détriment de l'apparition de nouveaux artefacts que nous appellerons les artefacts de correction.

Problématique Ces deux chapitres nous permettront d'établir la problématique de notre travail, à savoir comment supprimer des artefacts de compression sans faire apparaître d'artefacts de correction. Pour répondre à cette problématique et aux objectifs, nous proposons un travail qui se divise en deux parties.

Tout d'abord, une partie sera destinée à l'estimation de la qualité pour pouvoir évaluer quantitativement la gêne occasionnée par ces artefacts de compression et de correction.

Deuxième partie

Chapitre 3 Après avoir introduit la notion de qualité, nous verrons qu'elle peut être évaluée soit de manière subjective par des observateurs donnant des jugements en suivant un protocole bien précis, soit de manière objective par le développement de mesures automatiques appelées métriques objectives.

Chapitre 4 A partir de ces mesures de qualité, nous ciblerons la gêne occasionnée par les artefacts de compression en étudiant les méthodes existantes de la littérature. Mettant en avant leurs limitations, nous proposerons une métrique reprenant le coeur de ces méthodes et améliorant certaines caractéristiques.

Chapitre 5 Puis, pour compléter l'estimation de la qualité, nous étudierons les méthodes proposées dans la littérature pour évaluer les artefacts de correction. Cette étude nous permettra d'appréhender un nouveau principe que nous développerons dans une métrique.

Au terme de cette partie, nous serons capables par rapport à la problématique d'estimer la gêne visuelle occasionnée par les artefacts de correction et les artefacts de compression.

Nous consacrerons alors une dernière partie ciblée vers l'élaboration de méthodes aboutissant à un bon compromis entre le niveau de visibilité des artefacts de compression et des artefacts de correction.

Troisième partie

Chapitre 6 Pour ce faire, nous proposerons tout d'abord une solution pour chaque limitation que nous aurons mise en avant sur les corrections existantes. Ces méthodes de correction s'adapteront aux contraintes que nous avons évoquées et seront guidées par les métriques développées.

Chapitre 7 Enfin, nous finaliserons ce travail en proposant une solution à la problématique au travers d'un algorithme d'amélioration de la qualité des images compressées. L'algorithme reprendra les solutions proposées pour chaque limitation que nous aurons décrites dans le chapitre précédent. L'efficacité de cet algorithme sera mis en avant par l'estimation de la visibilité des artefacts de compression et de correction.

Un glossaire permettra de se rapporter rapidement aux différentes définitions et sigles que nous aurons pu définir dans un chapitre et réutiliser dans un autre.

Première partie

Analyse du sujet et Problématique

Chapitre 1

La Compression Numérique : Rôle, Principe et Conséquences

Nous débuterons ce chapitre par un historique sur les normes. Cela permettra de situer le MPEG-2 parmi les autres normes. Puis, nous étudierons plus précisément cette norme MPEG-2 en expliquant quels principes de compression sont appliqués sur un flux vidéo. Enfin, nous verrons quels types d'artefacts peuvent être causés par la combinaison de ces principes de compression à une limitation de la bande passante.

1 Les normes de compression

Les normes, spécifiques à des applications, sont définies par des groupes de travail appartenant à des organismes. Nous commencerons par définir ces organismes. Puis nous verrons l'évolution des normes au cours du temps.

1.1 Les organismes de normalisation

Au niveau international, les deux organismes les plus actifs pour la normalisation et les systèmes de compression vidéo sont l'ITU-T et l'ISO/IEC.

1.1.1 L'ITU-T : International Telecommunication Union, Standardization

Fondée en 1865, l'ITU ou l'institution spécialisée des nations unies pour les technologies de l'information et de la communication a contribué à aider le monde à communiquer en œuvrant dans 3 secteurs principaux : les radiocommunications, la normalisation et le développement des télécommunications. Le secteur de normalisation de l'ITU plus communément appelé l'ITU-T a été créé en 1993 dans le but d'établir des normes associées à la télécommunication.

1.1.2 L'ISO : International Organization of Standardization

L'ISO est un organisme de normalisation international composé de représentants d'organisations nationales de normalisation de 157 pays. Cette organisation créée en 1947 a pour but de produire des normes internationales dans les domaines industriels et commerciaux appelées normes ISO. L'ISO est le premier producteur mondial de Normes Internationales. L'ISO travaille conjointement avec la CEI, Commission Electrotechnique Internationale et le rôle de l'ISO/CEI est de fournir aux personnes concernées de l'information sur la normalisation et les normes.

1.1.3 Les normes

Au sein de ces deux organismes, des groupes de travail définissent des normes et chacune d'entre elles contient un ensemble de standards ou fascicules dédiés au flux système, au codage vidéo, ou encore à l'audio. Dans le cadre de ce travail, nous nous concentrerons plus précisément sur la partie codage vidéo de ces normes.

1.2 Les normes spécifiques aux images fixes

Le principe fondamental de la vidéo (succession d'images) est que le Système Visuel Humain ou SVH a la faculté de retenir pendant un certain temps (de l'ordre du dixième de seconde) toute image imprimée sur la rétine. Il suffit donc de faire défiler un nombre suffisant d'images par seconde, pour que le SVH ne se rende pas compte qu'il s'agit d'images distinctes. Pour ces raisons, il semble logique que les normes de compression vidéos fassent appel à des méthodes issues des normes de compression d'images fixes et c'est pourquoi nous commencerons par évoquer ces normes.

Le groupe JPEG La normalisation des standards de compression d'images fixes est éditée depuis à la fin des années 80 par un comité d'experts nommé Joint Photographic Experts Group. Ce comité d'experts appartient aux deux groupes de normalisation ISO/IEC et ITU-T.

La norme JPEG La première norme définie par ce groupe d'experts porte son nom. La norme JPEG ou JPG liée à la compression d'images fixes est apparue en 1989 et a été conçue à l'origine pour le monde de l'impression et de la photocomposition. JPEG accepte n'importe quelle définition d'image et compresse les données en exploitant les similarités à l'intérieur de chacune d'elle. Pour cela, l'image est découpée en bloc et cette étude se fait dans chacun des blocs.

La norme JPEG2000 En vue d'obtenir de meilleur taux de compression, le groupe JPEG a proposé au travers de la norme JPEG2000 en 2000 un système de compression totalement différent de celui proposé dans la norme JPEG. Cette technologie de compression utilise une transformation en ondelettes (méthode d'analyse mathématique du signal) qui offre une plus grande finesse au niveau de l'analyse du signal et permet de mieux s'adapter aux propriétés locales de l'image. De la même façon que pour le format JPEG, JPEG2000 est capable de travailler avec ou sans pertes.

1.3 De l'image à la vidéo

La vidéo étant une succession d'images, des principes de compression vidéo se sont développés à partir des techniques de compression d'images.

En effet, le Motion JPEG noté MJPEG ou M-JPEG consiste à appliquer successivement des systèmes JPEG capables de compresser ou décompresser en temps réel 25 ou 30 images par seconde. Etant donné que le M-JPEG code séparément chaque image de la séquence, il permet d'accéder aléatoirement à n'importe quelle partie d'une vidéo ce qui le rend utilisable dans les studios de montage numérique. Cependant, ces modifications permettant de passer du format JPEG au format MJPEG n'ont jamais été normalisées. Seul le format d'enregistrement DV (Digital Video) utilisant les principes du MJPEG a été normalisé.

De la même façon, le Motion JPEG 2000 ou MJPEG200 représente l'évolution de la compression JPEG2000 au format vidéo. En d'autres termes, une vidéo MJPEG2000 est une simple concaténation d'images au format JPEG 2000 légèrement modifié.

1.4 Les normes spécifiques à la vidéo

Même si ces normes ont la faculté d'être utilisées pour les images fixes et pour les vidéos, seule la redondance spatiale c'est à dire les similarités au sein d'une même image est prise en compte dans les systèmes de compression. De manière à développer et améliorer les taux de compression, l'idée est de tenir compte aussi de la redondance temporelle, c'est-à-dire des similarités trouvées au sein d'une succession d'images constituant une vidéo.

1.4.1 Les normes h26x

Le groupe VCEG Dans le but de définir des recommandations techniques destinées aux applications de visiophonie et de visioconférence, le groupe vidéo de l'ITU-T : le VCEG (Video Coding Experts Group) a défini plusieurs normes.

La norme H261 Première norme vidéo approuvée en 1993, H.261 [ITU, 1993] vise les applications de visiophonie pour le réseau RNIS à des débits multiples de 64 kbit/s. Les formats d'images traités sont le QCIF(144x176 pixels) et le CIF(288x352 pixels). De manière à optimiser

la compression, ce type de codage utilise bien les similarités à l'intérieur des images et entre les images successives d'une vidéo.

La norme H263 La norme de codage vidéo H263 [ITU, 1996] a ensuite été créée pour la communication vidéo à très bas débit dont la première version fut adoptée en 1995. Elle vise les applications de visiophonie et de visioconférence sur les réseaux RTC et RNIS. Cette norme repose sur les principes mis en place par la recommandation H.261. Les formats d'images sont des multiples et des sous-multiples du CIF (352x288 pixels). D'autres versions de cette norme H263, appelée H263+ [ITU, 1998] et H263++ [ITU, 2000] sont apparues respectivement en 1998 et en 2000 et proposent des options supplémentaires qui améliorent la qualité et la robustesse aux erreurs.

1.4.2 Les normes MPEG

Le groupe MPEG En parallèle à ces travaux ciblés sur la visioconférence, les travaux techniques de l'ISO/IEC sont menés au sein du groupe MPEG (Motion Picture Experts Group). Créé en 1988, le groupe MPEG compte environ 300 membres issus de diverses industries et universités et a défini les standards MPEG-1, MPEG-2, MPEG-4 pour des applications aussi variées que la télévision ou le multimédia.

La norme MPEG-1 Normalisée en 1992, MPEG-1 [ISO/IEC, 1993] est une norme de compression des images animées à faible résolution, destinée aux applications multimédias. Elle se concentre sur le taux de compression plutôt que sur la qualité d'image et cible une qualité équivalente au VHS (format SIF ou CIF) à un débit de 1.5 Mbit/s. Cette norme a été construite sur la base de H.261 dont elle reprend les principes et apporte des améliorations. MPEG-1 n'est plus utilisé aujourd'hui si ce n'est en compression du son (la partie audio du standard) avec le format MP3 pour le stockage de la musique.

La norme MPEG-2 La norme MPEG-2 [ISO/IEC, 1994] a été définie en 1996 pour les applications liées à la télévision numérique, que ce soit au niveau professionnel (production audiovisuelle par exemple) qu'au niveau du grand public (diffusion vers les postes TV). Elle reprend les principes de MPEG-1 en ajoutant les outils indispensables pour les applications télévisuelles comme par exemple la gestion des formats entrelacés. La résolution de l'image d'entrée s'étend de la définition standard (720x576) à la haute définition (1920x1080). Comme nous l'avons mentionné dans l'introduction, ce standard a été adopté par le consortium DVB (Digital Video Broadcasting) pour les services de TV numérique par voie hertzienne terrestre (DVB-T) et satellite (DVB-S). Il est également utilisé comme format de codage du DVD (Digital Video Disc).

La norme MPEG-4 part2 Le premier successeur du MPEG-2 fut le MPEG-4 Part2. Cette nouvelle norme a permis quelques améliorations par rapport au MPEG-2 mais sans procurer un gain énorme. C'est de cette norme que dérivent les formats couramment rencontrés tels que le DivX, le Xvid et le NeroDigital. Pour éviter la confusion avec la norme suivante, nous précisons que nous parlons de la partie 2 de la norme MPEG-4 et l'appellerons MPEG-4 part2.

1.4.3 La fusion

Le groupe VCEG Le groupe VCEG de l'ITU-T a travaillé sur des applications de visioconférence et le groupe MPEG de l'ISO/IEC sur des applications aussi variées que la télévision ou le multimédia. On a pu remarquer qu'ils exploitent tous les deux les similarités à l'intérieur des images qui composent la vidéo et entre ces images. Soucieuses de mettre au point un système de compression plus efficace, ces deux instances de normalisation ont fusionné donnant naissance au groupe JVT (Joint Video Team) chargé de développer le système de codage AVC (Advanced Video Coding).

La norme H264, MPEG-4 AVC ou MPEG-4 Part10 La première version de la norme a été approuvée en 2003. Chaque organisme ayant tenu à lui attribuer un nom, cette norme porte la double dénomination MPEG-4 part10 AVC et H264. Elle utilise également les similarités au sein d'une même image et entre plusieurs images mais cette nouvelle norme [Richardson, 2003] présente des avancées notoires se traduisant par une efficacité de compression globalement deux fois supérieure à celle obtenue par la version MPEG-2 pour une qualité visuelle équivalente. Ce standard apparaît aujourd'hui dans le monde de la télévision numérique et des supports de stockage car il permet notamment une diffusion plus importante de la haute définition.

Le SVC Depuis 2004, le groupe JVT développe le SVC [Balter et al., 2006] ou Scalable Video Codec qui est l'extension de la norme MPEG-4 AVC à laquelle sont ajoutés des outils de scalabilité temporelle, spatiale et en qualité. Le standard est actuellement en cours de finalisation.

Pour l'avenir, les groupes de normalisation travaillent à l'apparition d'une norme intégrant les aspects de multi-vues et de profondeur.

1.4.4 L'AVS ou le standard audio video chinois

Il paraît important de citer la norme AVS (Audio Video coding Standard workgroup of China) qui a été normalisé par le standard Chinois et dont les principes reprennent les mêmes que ceux inclus dans la norme H264. Ce standard n'apporte rien de nouveau en terme de compression mais est en pleine expansion en Chine et risque donc d'avoir un rôle non négligeable dans l'avenir.

1.5 Un bilan sur les normes et les applications

Le tableau 1.1 résume le bref descriptif que nous venons de faire sur les normes de compression. Les dates restent approximatives car il y a souvent une légère variation entre le moment de la normalisation et la version améliorée de la norme réellement utilisée sous l'appellation donnée.

TAB. 1.1 – Les normes et leurs applications

Norme	Organisme	Groupe	Date	Application
JPEG	ITU-T - ISO	JPEG	1990	Images Fixes
H261	ITU-T	VCEG	1993	Visiophonie
H263	ITU-T	VCEG	1996	Visioconférence
H263++	ITU-T	VCEG	2000	Réseaux IP et mobiles
MPEG-1	ISO/IEC	MPEG	1992	Multimédia qualité VHS
MPEG-2	ISO/IEC	MPEG	1996	Vidéo Broadcast
MPEG-4 part2	ISO/IEC	MPEG	2000	Multimédia
MPEG-4 AVC/ H264	ITU-T et ISO/IEC	JVT	2003	Multimédia Haute Qualité

1.6 Discussion

Malgré l'efficacité des principes proposés dans la norme MPEG-4 AVC, c'est encore la norme MPEG-2 qui est majoritairement utilisée dans le domaine de la télévision numérique pour les raisons suivantes :

- Les encodeurs MPEG-2, aussi bien en définition standard qu'en haute définition sont parfaitement au point, au terme de plus de 15 ans de *R&D*.
- Le codage de compression MPEG-4 AVC est incontestablement plus performant que le MPEG-2. Cependant, cette amélioration est acquise au prix d'une complexité nettement supérieure et les performances de la microélectronique ne rendent parfois possible qu'une implémentation partielle de ces algorithmes sophistiqués.
- Enfin, un passage définitif à l'encodage MPEG-4 AVC nécessiterait le changement ou la mise à jour de tous les décodeurs MPEG-2 du marché qui ne sont pas compatibles avec la norme MPEG-4 AVC.

Cet historique permet de situer d'un point de vue technologique mais aussi chronologique la norme MPEG-2. En effet, l'intérêt était de montrer l'évolution et donc le lien entre ces différentes normes et de mettre en avant le fait qu'une même base est toujours utilisée pour compresser les données. Cette base consiste non seulement à étudier la redondance d'information à l'intérieur d'une même image composant la vidéo mais aussi à étudier la redondance d'information entre les images d'une même vidéo. Ce sont ces principes que nous allons développer dans la section suivante en nous intéressant plus particulièrement à la norme MPEG-2. Les lecteurs intéressés par le contenu des autres normes peuvent se référer aux documents officiels et aux documents de la littérature [Ebrahimi et Pereira, 2002],[Richardson, 2003], [Watkinson, 2004].

2 Les principes de la compression vidéo

La compression spatiale s'effectue en étudiant les similarités ou redondances dans une même image. La compression temporelle s'effectue en étudiant les similarités entre plusieurs images successives. Après avoir défini le contenu d'un flux vidéo, nous verrons quels sont les outils de codage et de compression utilisés pour ces deux types de compression.

2.1 Le flux vidéo

Les données d'un flux vidéo sont hiérarchisées selon une manière bien précise schématisée sur la figure 1.1. Nous allons décrire chacune de ces données.

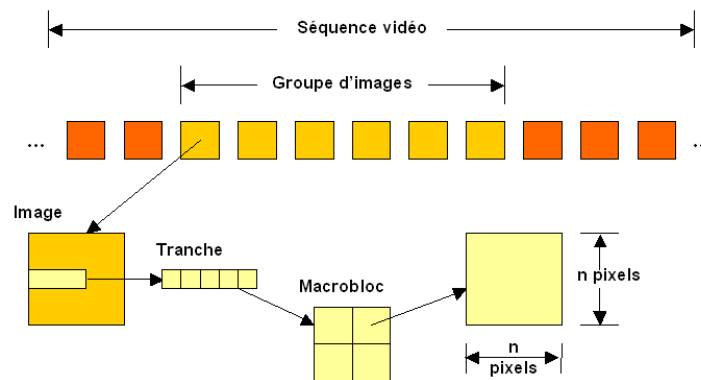


FIG. 1.1 – Hiérarchie dans le flux vidéo.

La séquence vidéo Elle commence par un code de début de séquence, contient un ou plusieurs groupes d'images et se termine par un code de fin de séquence.

Les groupes d'images ou GOP (Group of Pictures) Chaque groupe d'images regroupe un en-tête et une série d'une ou plusieurs images permettant d'y accéder de façon aléatoire.

L'image et sa représentation L'image numérique est composée de pixels. Les images de ces GOP seront représentées dans un espace qui va permettre, avant les techniques de compression, de réduire l'information. Ce type d'espace est un espace à trois dimensions : la composante d'intensité plus fréquemment appelée par sa traduction anglaise "luminance" et deux composantes chromatiques plus fréquemment appelées par le terme anglais "chrominance". Les professionnels de la télévision numérique utilisent l'espace YCbCr (Luminance Y, différence chromatique bleue-jaune Cb et différence chromatique rouge-verte Cr). Sur cet espace, il est alors possible de réduire le nombre des échantillons Cb et Cr à transmettre. Ces choix dépendent des applications et sont exprimés au travers de formats récapitulés dans le tableau 1.2 sur une base d'une image de taille 720x576 pixels (pour une numérisation en 8 bits). Par exemple, pour le format 4 :2 :0, il y a un élément de luminance (Y) pour chaque pixel, et des éléments de chrominance (U et V) un pixel sur deux, et une ligne sur deux en alternance. La résolution de la chrominance est divisée par 2 (360x288).

Remarque : L'utilisation de cet espace a un autre intérêt : il permet de rester compatible aux postes de télévision noir et blanc sur lesquels ne sera envoyée que l'information de luminance.

TAB. 1.2 – Sous-échantillonnage des composantes Cb et Cr

Format	Luminance Y	Chrominance Cr et Cb	Echantillonnage	Utilisation
4 :4 :4	720x576	720x576	24bits/pixel	Aucune
4 :2 :2	720x576	360x576	16bits/pixel	DVD, Montage
4 :2 :0	720x576	360x288	12bits/pixel	Télévision numérique

La tranche La hiérarchisation des données continue par le découpage de chaque image en tranches. Une tranche est constituée de un ou plusieurs macroblocs adjacents ordonnés de gauche à droite puis de haut en bas. Ce sont des éléments importants pour la gestion des erreurs. Si le flux de données contient une erreur, le décodeur peut sauter la tranche et passer au début de la suivante directement.

Le macrobloc ou MB Le macrobloc est une matrice de dimension 2. Dans le cas de la luminance, ce MB est constitué de 4 blocs. Dans le cas de la chrominance, la taille du MB dépend du sous-échantillonnage effectué. Par exemple, dans le cas du 4 :2 :0 représenté sur la figure 1.2 l'information de chrominance étant choisie un pixel sur deux et une ligne sur deux, la taille du MB final sera de 8 pixels par 8 pixels.

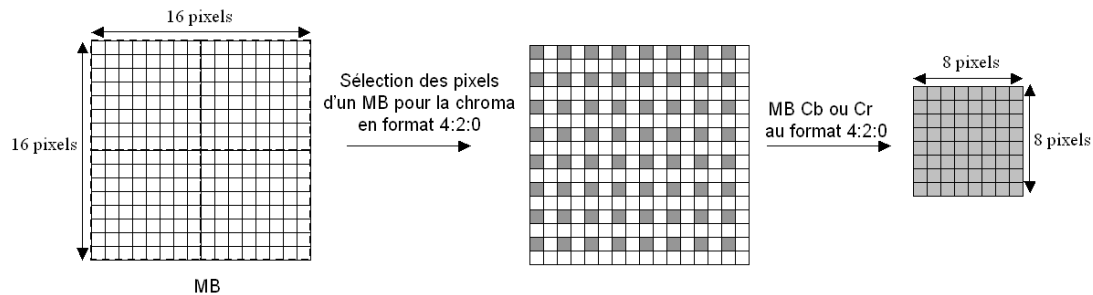


FIG. 1.2 – Sélection des pixels de chrominance dans le cas d'un format 4 :2 :0.

Le bloc Enfin, le bloc est un ensemble de valeurs de luminance ou de chrominance constitué de n lignes et de m colonnes. Dans le cas d'un MB en luminance codé en MPEG-2, $n = m = 8$. Chaque bloc est représenté par un tableau de 64 nombres entiers (dans le cas d'un bloc 8x8 pixels) compris entre 0 et 255 pour un codage sur 8 bits. C'est au sein de ce bloc que les techniques de compression vont être appliquées.

2.2 La compression spatiale

La compression spatiale exploite les redondances spatiales, subjectives et statistiques à l'intérieur d'un bloc.

2.2.1 La redondance spatiale

L'étude de la redondance spatiale consiste à mettre en avant les redondances existantes au sein d'une même image appartenant au flux vidéo et se traduit par deux étapes significatives : la décomposition en blocs et le passage dans le domaine fréquentiel.

La décomposition en blocs Quelques soient les transformations effectuées au cours de la compression, il paraît compliqué de les effectuer d'un trait sur l'ensemble de l'image car cela implique la manipulation de quantités importantes de données et donc nécessite des temps de calcul prohibitifs. La première étape, illustrée sur la figure 1.3 consiste donc à découper l'image en macroblocs eux mêmes re-découpés en blocs.

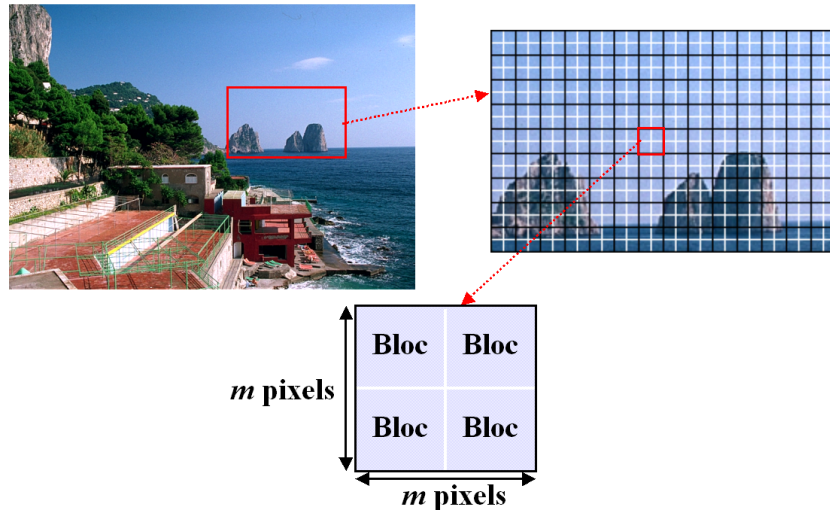


FIG. 1.3 – Découpage d'une image en Macroblocs et en Blocs.

Le passage dans le domaine fréquentiel Le système visuel humain se satisfait d'une résolution assez faible pour les détails d'une image : cette propriété (que nous expliquerons au chapitre 3 section 4.1.1) est la base du principe de compression spatiale. Pour la mettre en œuvre, elle nécessite une transformation vers le domaine fréquentiel pour mettre clairement en évidence les informations à modifier. Ce passage s'effectue par l'intermédiaire d'une transformée mathématique linéaire et bidimensionnelle appelée Transformée en Cosinus Discrète (DCT). Elle permet de décrire chaque bloc en une carte de fréquences et d'amplitudes plutôt qu'en pixels. Pour un bloc de taille $N \times N$, la transformée DCT s'exprime mathématiquement par la formule suivante :

$$DCT(i, j) = \frac{2}{N} C(i) C(j) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} pixel(i, j) \cos \left[\frac{(2x+1)iH}{2N} \right] \cos \left[\frac{(2y+1)jH}{2N} \right] \quad (1.1)$$

où C est une constante qui vaut :

$$C = \begin{cases} \frac{1}{\sqrt{2}} & \text{pour } x = 0 \\ 1 & \text{pour } x > 0 \end{cases} \quad (1.2)$$

Cette DCT va permettre de décrire les composantes du bloc en fréquences spatiales représentant des motifs géométriques périodiques caractérisés par leur taille et leur amplitude. La figure 1.4 donne l'exemple d'une bibliothèque de motifs que l'on peut obtenir avec un bloc DCT de taille 8×8 . Les fréquences spatiales horizontales augmentent de la gauche vers la droite (les détails horizontaux deviennent de plus en plus fins) et les fréquences spatiales verticales croissent de haut en bas (les détails verticaux deviennent de plus en plus fins).

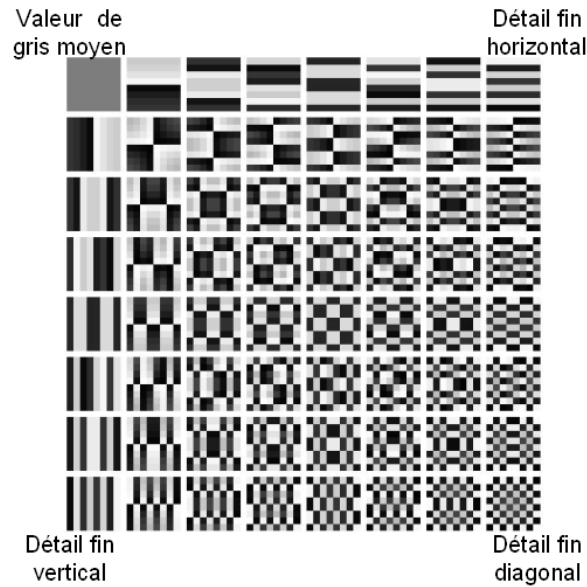


FIG. 1.4 – Catalogue des motifs d'un bloc DCT 8x8.

La conversion de l'information Ainsi, les 64 nombres représentant les valeurs de luminance (ou de chrominance) des pixels de chaque bloc sont convertis par la DCT en autant de coefficients d'amplitude appliqués aux composantes fréquentielles. Ces 64 composantes fréquentielles permettent donc de caractériser tous les motifs basiques (figure 1.4) qu'il est possible d'obtenir avec les 64 pixels du bloc. Les coefficients d'amplitude indiquent dans quelle proportion chaque fréquence spatiale est présente dans le bloc considéré. La DCT fournit ainsi une mesure directe de la quantité de détails présente dans ce bloc.

L'interprétation des coefficients la première case en haut à gauche du bloc DCT contient le coefficient le plus élevé, le coefficient DC, valeur proportionnelle à la moyenne du bloc (fréquence nulle=composante nulle). Plus on s'éloigne de cette case en se dirigeant vers celle située en bas à droite, plus les motifs rencontrés sont fins et représentent les coefficients AC. Sur la figure 1.5, nous avons représenté deux exemples de blocs de contenus différents nous permettant d'interpréter leurs coefficients DCT. Le premier exemple montre que si la zone d'image analysée est uniforme, seule la première case du bloc est remplie. Dans ce cas particulier, la DCT permet de décrire les 64 pixels d'un bloc avec une seule valeur. Le second exemple montre sur un bloc contenant de l'information que les valeurs importantes sont concentrés dans la partie supérieure de la diagonale du bloc. La DCT permet de concentrer les valeurs significatives sur une zone restreinte. Notons bien que cette transformée purement mathématique préserve l'intégralité de l'information du bloc et n'introduit aucune perte (si on considère comme négligeable les pertes dues au passage dans le domaine des entiers des coefficients DCT) : c'est une transformation réversible.

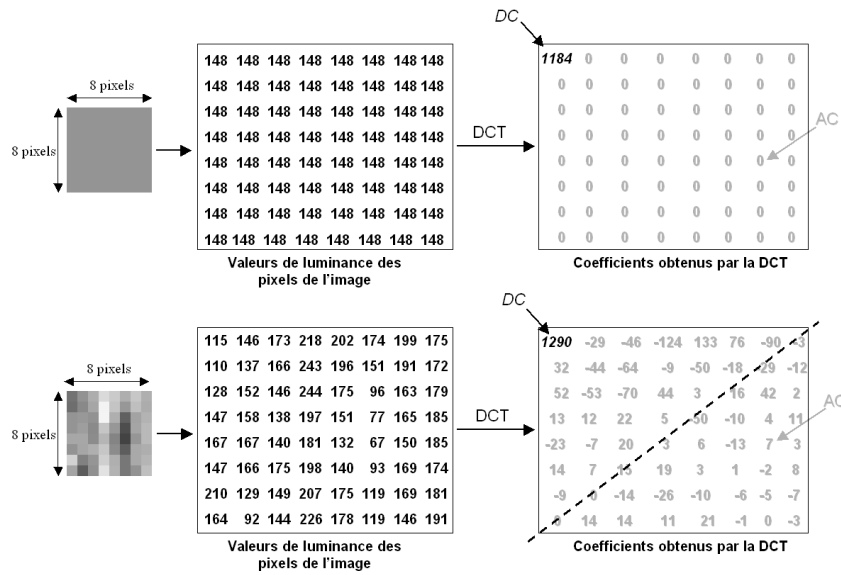


FIG. 1.5 – Exemples de passage des valeurs des pixels dans le domaine fréquentiel avec la DCT pour un bloc homogène et un bloc non-homogène.

2.2.2 La redondance subjective

Comme le système visuel humain accorde moins d'importance aux amplitudes des détails fins de l'image qu'aux plages uniformes, les hautes fréquences spatiales vont pouvoir être codées avec une fidélité moins grande que les basses fréquences et donc être sous quantifiées. Certains coefficients inférieurs à un certain seuil seront même éliminés.

Les tables de quantification Cette phase est effectuée en utilisant différentes tables de quantification employées pour la luminance et la chrominance qui déterminent pour chacun des coefficients DCT par quelle valeur il sera quantifié. Cette étape est représentée sur la figure 1.6. Les coefficients de la matrice de quantification sont choisis en fonction de la capacité du SVH à détecter les motifs par leur changement de contraste. Effectuée dans le but de supprimer des données non visibles à l'œil, cette redondance est appelée redondance subjective.

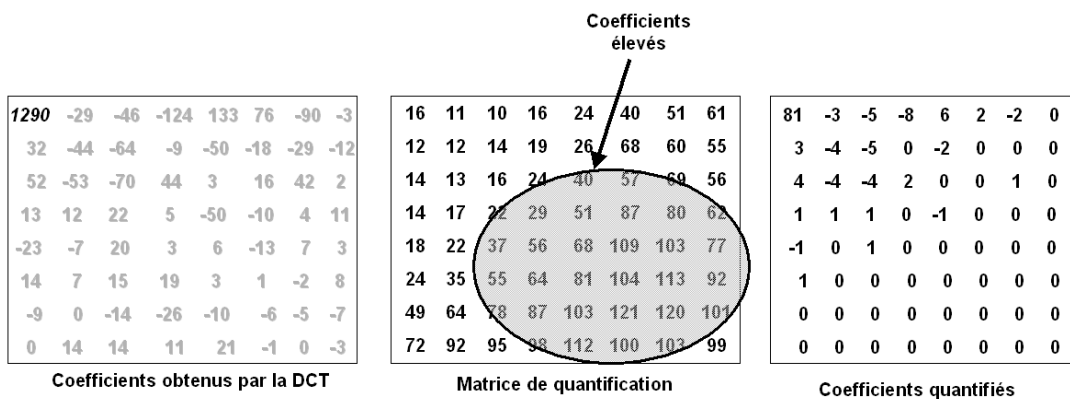


FIG. 1.6 – Effet de la quantification sur les coefficients DCT.

Les pertes d'informations Plus les coefficients sont quantifiés, moins il restera d'informations à coder. Néanmoins, c'est dans cette phase de quantification que sont introduites les pertes d'informations et que des distorsions peuvent apparaître. Car même si ces tables sont élaborées en tenant compte des caractéristiques du système visuel humain, la limitation de la bande passante peut être responsable de l'augmentation des valeurs de quantification et donc d'une perte d'information trop importante pour pouvoir coder correctement les caractéristiques du bloc considéré.

2.2.3 La redondance statistique

Après avoir exploité la redondance spatiale et la redondance subjective, les données peuvent encore être compressées en utilisant la redondance statistique. Ces méthodes sont employées depuis des années au travers d'autres applications comme par exemple le codage MORSE et permettent de coder intelligemment l'information en fonction de leur importance et leur occurrence.

Le balayage en zig zag Tout d'abord, la matrice obtenue après quantification est balayée en zig zag comme l'illustre la figure 1.7. Les coefficients sont réarrangés sous la forme d'un vecteur dans lequel les premières composantes représentent les basses fréquences et les dernières les hautes fréquences. L'objectif est de concentrer les zéros en fin de vecteur.

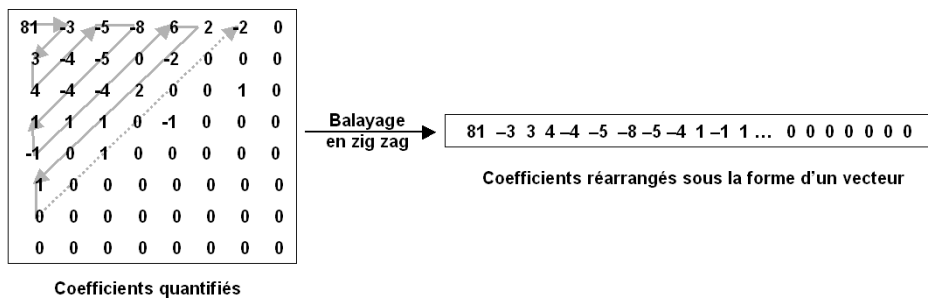


FIG. 1.7 – Balayage en zig zag de la matrice DCT quantifiée.

Le codage à longueur variable Le vecteur ainsi formé contient des suites plus ou moins longues de coefficients de valeurs identiques, souvent nulles. Plutôt que de coder systématiquement chaque coefficient à chaque fois qu'il apparaît, le codage à longueur variable (ou Run Length Coding) formate le vecteur sous la forme de paires de données. La première donnée indique la valeur du coefficient, tandis que la seconde signale le nombre de fois qu'il se répète. Ce codage est plus communément appelé le RLC (Run Length Coding).

Le codage entropique Une analyse statistique de l'occurrence des coefficients est alors effectuée pour procéder à un codage entropique. Son principe consiste à attribuer les codes les plus courts aux coefficients statistiquement les plus fréquents, et à réserver les codes longs aux coefficients dont la probabilité d'occurrence est faible.

Remarque : le codage entropique et le codage à longueur variable sont tous deux des algorithmes sans pertes, qui permettent à eux seuls un taux de compression de l'ordre 2 :1.

2.2.4 Synthèse de la compression spatiale

Pour résumer, les étapes utilisées pour la compression spatiale sont les suivantes :

- Découpage de l'image en MB eux mêmes subdivisés en blocs.
- Passage au domaine fréquentiel par le calcul de la DCT.
- Quantification des coefficients DCT.
- Balayage en zig zag transformant la matrice en vecteur de données.
- Codage à longueur variable et codage entropique pour exploiter les redondances statistiques du vecteur de données.

Si on considère comme négligeable les pertes dues au passage en entier des coefficients DCT lors du passage dans le domaine fréquentiel, les seules pertes possibles lors de la compression spatiale sont introduites par l'étape de quantification.

2.3 La compression temporelle

La redondance temporelle va permettre d'exploiter les redondances entre les images en ne transmettant que les différences entre deux images consécutives. La compression temporelle s'applique aux groupes d'image de la séquence vidéo ou GOP (Group of Pictures) représentées sur la figure 1.1.

2.3.1 Les GOP

Les GOP se composent d'une combinaison de trois types d'images : les images Intra ou images I, les images Prédites ou images P et les images Bidirectionnelles ou images B. La figure 1.8 représente un exemple de structure de GOP que nous allons détailler.

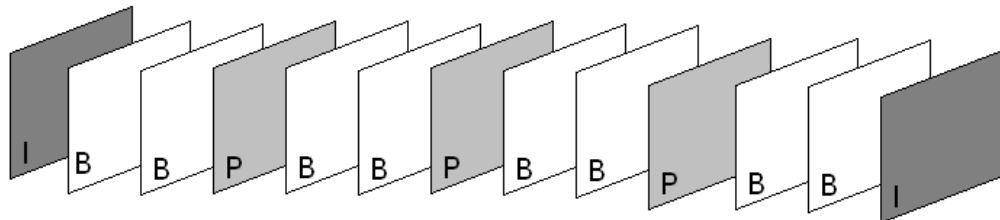


FIG. 1.8 – Exemple d'organisation d'un GOP de 12 images.

L'image I Elle est codée en mode intra-image. Le terme intra-image signifie que seules les données à l'intérieur de l'image sont exploitées. L'image I est donc codée avec les principes de compression spatiale évoqués précédemment.

Remarque : l'image I est donc entièrement décrite par elle-même, sans aucune références à d'autres, et contient tous les éléments nécessaires à sa reconstruction. Elle constitue de ce fait le point d'accès pour le décodage. La fréquence d'occurrence des images I dans une séquence vidéo conditionne la précision de l'accès aléatoire.

L'image P Elle est prédite à partir d'une image passée I ou P. Elle est codée uniquement à l'aide de vecteurs de mouvement indiquant les déplacements de ses éléments par rapport à l'image de référence.

Remarque : une image P est typiquement trois fois moins volumineuse qu'une image I, mais peut

transmettre des erreurs car elle sert également de référence pour les images B que nous allons définir. Il faut attendre l'arrivée d'une image I pour tout remettre à plat et relancer un nouveau processus de prédiction.

L'image B est construite à l'aide de vecteurs de mouvement, par interpolation bidirectionnelle entre les images passées ou futures I ou P voisines.

Remarque : l'image B offre le taux de compression le plus élevé. Une image B est typiquement six fois moins volumineuse qu'une image I. C'est grâce aux images B que l'on peut faire chuter considérablement le débit d'un flux vidéo.

A titre d'exemple, la figure 1.9 illustre la quantité de données à coder pour une suite d'images I, P et B d'une même séquence. Toutes les informations sont codées pour une image I tandis que seules les informations en mouvement (eau de la piscine et nageuses) sont codées pour les images P et B.



FIG. 1.9 – De gauche à droite : Informations à coder pour une image I, une image P et une image B.

2.3.2 La compensation de mouvement

C'est la technique de compensation de mouvement qui va permettre d'obtenir ces images P et B.

L'estimation de mouvement La première étape est l'estimation de mouvement. Son rôle est de construire une image de prédiction à partir d'une image précédente et d'informations relatives aux déplacements subis par ces composantes. L'estimation de mouvement n'est pas effectuée sur chaque point de l'image ni même par les blocs de la DCT mais sur des macroblocs. Si entre une image actuelle et celle qui la précède, deux macroblocs semblables sont repérés, mais à des emplacements légèrement différents, il suffit de transmettre une seule fois ce macrobloc et d'indiquer quelle est sa nouvelle position sur l'image actuelle. Cette position est définie par un vecteur de mouvement.

Le block matching La technique la plus répandue pour former le vecteur mouvement est celle du block matching, que l'on traduit par correspondance des blocs. L'estimateur de mouvement compare l'image d'entrée que nous considérons comme la nouvelle image t avec l'image précédente $t-1$ conservée en mémoire. Cette comparaison consiste à examiner un à un les macroblocs de l'image t afin de voir s'ils existaient sur l'image $t-1$.

La figure 1.10 illustre ce principe avec un macrobloc sur l'image t que l'on cherche à localiser sur l'image $t-1$. Une exploration est alors réalisée dans toutes les directions possibles à l'intérieur d'une fenêtre de recherche, afin d'identifier le macrobloc qui lui ressemble le plus. Lorsqu'il est repéré, la différence de position spatiale du macrobloc entre les deux images permet de déterminer les coordonnées du vecteur mouvement. Son amplitude représente la vitesse de déplacement et sa direction indique celle de la translation.

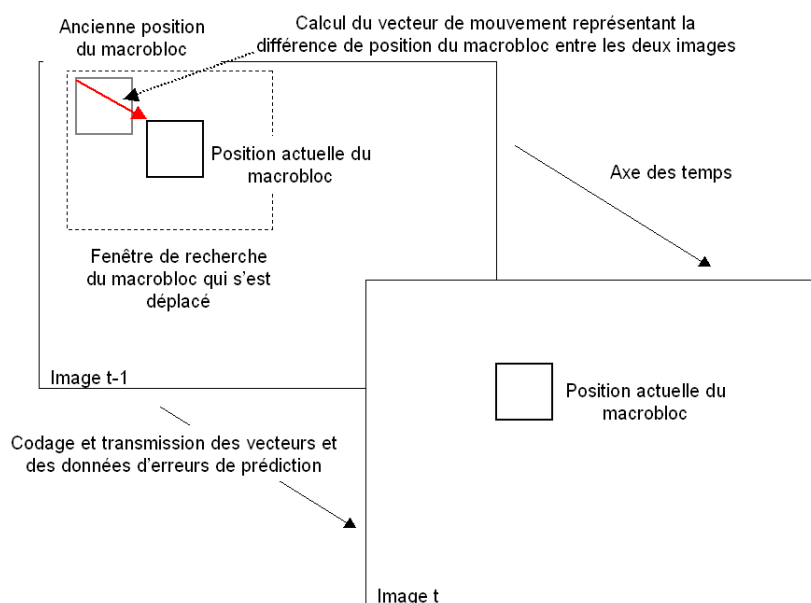


FIG. 1.10 – Principe de l'estimation de mouvement.

Le vecteur d'erreurs Toutes les correspondances étroites de MB sont combinées pour générer une image prédite. Cette image prédite est comparée avec la vraie image t afin de produire des données de différence visant à compenser les erreurs ou imprécisions de l'estimateur de mouvement. Les vecteurs mouvement et les données de différence sont transmis avec l'image $t-1$ (voir figure 1.10) et codés de la même façon qu'un bloc est codé spatialement (DCT, quantification et redondance statistique). Ils signaleront au décodeur comment il devra déplacer les MBs de cette image $t-1$ pour construire l'image t .

L'intérêt du codage bidirectionnel ou le rôle des images B La technique de compensation de mouvement avec les images P est très puissante mais ne résout pas tous les problèmes de la compression vidéo. Par exemple, considérons une scène où un personnage ouvre une porte, il n'y a aucun moyen de prédire les détails de la pièce qui se trouvent derrière la porte avec l'image précédente où la porte était fermée. L'intérêt du codage bidirectionnel est qu'il peut parfois être plus efficace de chercher les données dans l'image suivante que dans l'image précédente. L'inconvénient de ces images B est que le codeur décodeur (codec) MPEG est obligé de changer l'ordre des images dans le flux vidéo car le décodeur ne pourrait pas décodeur une image B sans avoir reçu l'image de référence si celle-ci se situe plus loin dans la séquence vidéo.

2.3.3 Synthèse de la compression temporelle

La compression temporelle est créée à partir des images I, des images P et B grâce aux cinq étapes de la compensation de mouvement :

- Recherche des macroblocs semblables entre une image t et l'image précédente $t-1$.
- Codage des vecteurs mouvement caractérisant les déplacements des macroblocs.
- Construction d'une image prédite en utilisant ces vecteurs de mouvement.
- Comparaison de cette image prédite avec la vraie nouvelle image pour générer des données d'erreurs de prédiction
- Codage et transmission des vecteurs et des données d'erreurs de prédiction

Au niveau de la compression temporelle, les pertes d'informations peuvent être introduites par une mauvaise estimation du vecteur de mouvement. Dans ce cas, cette erreur est minimisée grâce à la prédiction de l'erreur obtenue entre le MB prédit et le MB de référence. Néanmoins, cette erreur étant codée comme un bloc de données par la compression spatiale, elle est soumise aux mêmes possibilités de pertes dues à la compression spatiale (étape de quantification).

Cette section nous a permis d'analyser les méthodes utilisées pour compresser les données d'une vidéo numérique compressée en MPEG-2. Des précisions sur les techniques d'encodage et décodage sur cette norme sont expliquées dans les ouvrages [Symes, 2003] et [Bellaïche, 2003].

3 Les artefacts de compression

Nous allons voir les conséquences des pertes causées par la compression spatiale et la compression temporelle sur les images compressées décodées. Nous différencierons les distorsions spatiales, c'est à dire qui apparaissent sur le contenu d'une image des distorsions temporelles qui n'apparaissent seulement lorsque les images défilent. Cette partie a été inspirée de l'état de l'art des artefacts de compression fait par Yuen [Yuen et Wu, 1998] en 1998. Sauf précision, ces artefacts peuvent exister sur la composante de luminance comme sur les composantes de chrominance.

3.1 Les distorsions spatiales

3.1.1 L'effet de bloc

La cause L'image est découpée en MBs eux-mêmes subdivisés en blocs. La quantification est effectuée sur chacun des blocs traités indépendamment les uns des autres. Si cette quantification est trop forte, lors de la reconstruction de deux blocs adjacents, il est probable que la continuité de l'information présente dans l'image non compressée soit cassée au niveau des frontières de blocs.

La conséquence Une discontinuité de l'information apparaît alors au niveau de ces frontières de blocs et crée l'effet de bloc. La figure 1.11 illustre cet effet : l'effet de bloc apparaît à une fréquence de 8 pixels correspondant au découpage de l'image par blocs.

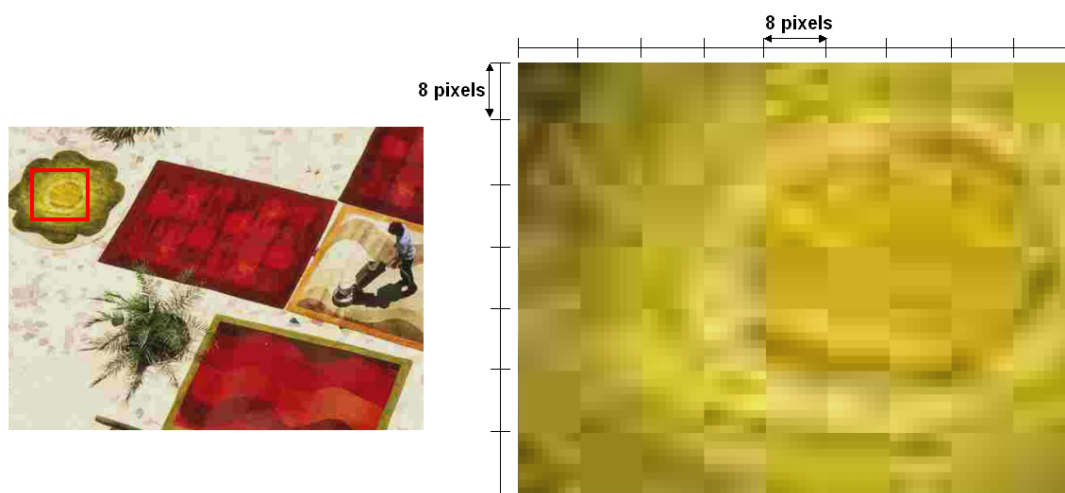


FIG. 1.11 – L'effet de bloc.

Les différents niveaux de l'effet de bloc La gêne occasionnée par l'effet de bloc va dépendre non seulement du taux de quantification mais aussi du contenu du bloc et de ces voisins. En effet, plus le bloc va contenir des variations d'informations, plus la frontière sera visible : cet effet est appelé l'effet de masquage. La figure 1.12 illustre cet effet de masquage : l'effet de bloc est beaucoup plus visible sur le ciel (zones avec peu de détails) que sur l'immeuble (zone texturée contenant des variations d'informations).

Remarque : Plus le taux de quantification des blocs sera élevé, plus leur contenu tendra vers une zone avec peu d'information : l'effet de masquage sera supprimé rendant l'effet de bloc plus visible.

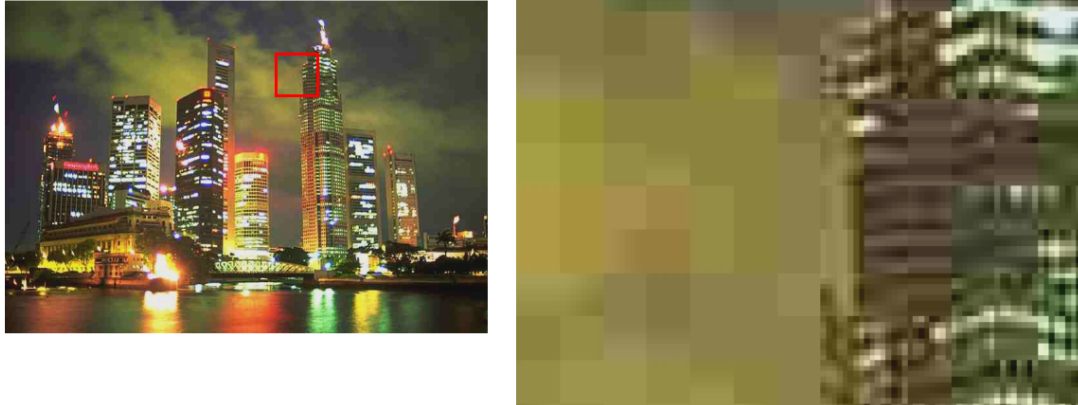


FIG. 1.12 – Effet de masquage : l'effet de bloc est plus visible dans les zones avec peu de détails que dans les zones contenant des variations d'informations.

3.1.2 L'effet de motif

La cause Après la DCT, le contenu d'un bloc est représenté par une combinaison de motifs. L'importance de chaque motif est donnée par les coefficients AC. Après la quantification, il est possible qu'il ne reste plus qu'un ou que très peu de coefficients AC.

La conséquence Le bloc résultant de la quantification représentera alors le motif correspondant au coefficient AC. La figure 1.13 montre un exemple de l'effet de motif : les motifs de l'image peuvent aisément être retrouvés dans la liste des motifs.

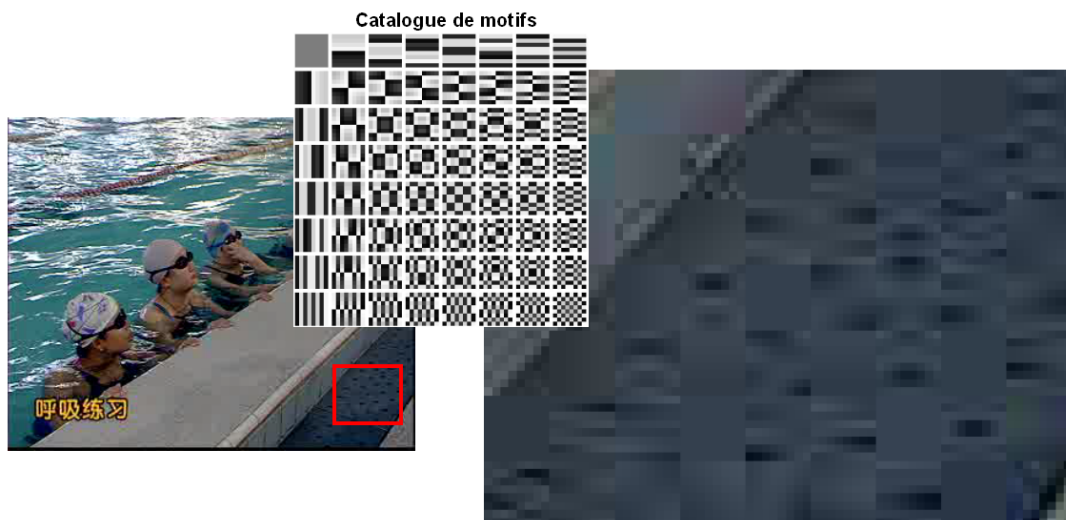


FIG. 1.13 – L'effet de motif.

Les différents niveaux de l'effet de motif Les coefficients de la matrice de quantification sont choisis en fonction de la capacité du SVH à détecter les motifs par leur changement de contraste. De plus, la position des coefficients est telle que le balayage en zig zag de cette matrice favorise les motifs basse fréquence aux motifs haute fréquence. L'impact visuel de l'effet de motif sera non seulement proportionnel à l'amplitude du coefficient AC mais aussi à sa position (dans l'ordre du zig zag) sur la matrice.

Remarque : L'apparition de ces motifs n'est pas forcément liée à l'isolement d'un coefficient AC mais peut être dû à la superposition de plusieurs motifs ayant la même orientation. Inversement, plus les motifs avec des orientations différentes seront superposés, moins on pourra reconnaître un motif et moins l'effet de motif sera perceptible.

3.1.3 Le flou

La cause L'étape de quantification privilégie les coefficients basses fréquences aux coefficients hautes fréquences. Après quantification, toutes les informations liées au détail sont donc atténuées ou supprimées.

La conséquence Même si l'image reste compréhensible dans son contenu, les hautes fréquences représentant les contours ou les textures sont atténuées et provoquent un effet de flou sur l'ensemble de l'image. Sur la figure 1.14, même si on comprend que ce sont 4 nageuses dans l'eau, l'image ne nous semble pas nette. En effet, les mains se confondent avec le bord de la piscine et les bonnets de bain sont dans la continuité de l'eau.



FIG. 1.14 – Le flou.

Les différents niveaux du flou Plus la quantification des blocs contenant des variations d'information sera forte, plus le flou sera important.

Remarque : Dans le cas d'une image naturelle, la quantification des blocs contenant très peu de détail fera tendre le contenu du bloc à une zone sans variation représentée uniquement avec le coefficient DC. Dans ce cas, l'effet de flou laissera place à un effet non naturel qui fera penser à une image de synthèse.

3.1.4 Les fausses couleurs

La cause Dans l'espace (YCbCr), l'information de couleur est donnée par les composantes Cb et Cr. On peut représenter l'information de la composante Cb sur un axe allant du bleu au jaune et l'information de la composante Cr sur un axe allant du vert au rouge. Sur la figure 1.15, nous avons décomposé une partie d'une image sous ses trois composantes. On comprend mieux alors comment la combinaison des trois composantes donne l'image finale. Après quantification, la valeur correspondant à une couleur sur un axe peut être modifiée par l'autre couleur de l'axe.



FIG. 1.15 – La décomposition d'une image avec les composantes Y, Cb et Cr.

La conséquence Lors de la réunification des trois composantes, une fausse couleur sera alors représentée. La figure 1.16 illustre ce phénomène et on remarque d'ailleurs que dans les zones à dominante jaune, la fausse couleur tend vers le bleu et dans les zones à dominantes vertes, la fausse couleur tend vers le rouge.

Les différents niveaux de fausses couleurs Plus la couleur obtenue après quantification sera loin sur l'axe des couleurs de la couleur initiale, plus la couleur sera "fausse" et donc visible. Ce phénomène est accentué par le sous-échantillonnage des composantes de couleur : par exemple, dans le format 4 :2 :0, si la quantification crée une fausse couleur, elle sera représentée sur 4 pixels et sera d'autant plus visible.

Remarque : La quantification ne tient pas compte de la sensibilité du SVH aux différentes couleurs. En effet, de part sa constitution, l'œil est plus sensible aux variations des couleurs vertes. On pourrait alors imaginer une quantification plus grossière de la composante Cb au profit de la composante Cr. On peut également appliquer ce raisonnement au sous-échantillonnage.



FIG. 1.16 – Les fausses couleurs.

3.1.5 L'effet d'escalier

La cause Pour représenter un contour droit mais non vertical ou horizontal, l'idéal serait d'utiliser une combinaison des motifs hautes fréquences située à la fin (dans l'ordre du zig zag) de la matrice DCT. Or, ces coefficients sont généralement quantifiés et le contour est représenté par une combinaison de basses fréquences horizontales ou verticales.

La conséquence Lorsqu'un contour droit est représenté par des blocs adjacents, le contour sera reconstruit comme une succession de contours horizontaux et verticaux. L'effet d'escalier est représenté sur la figure.

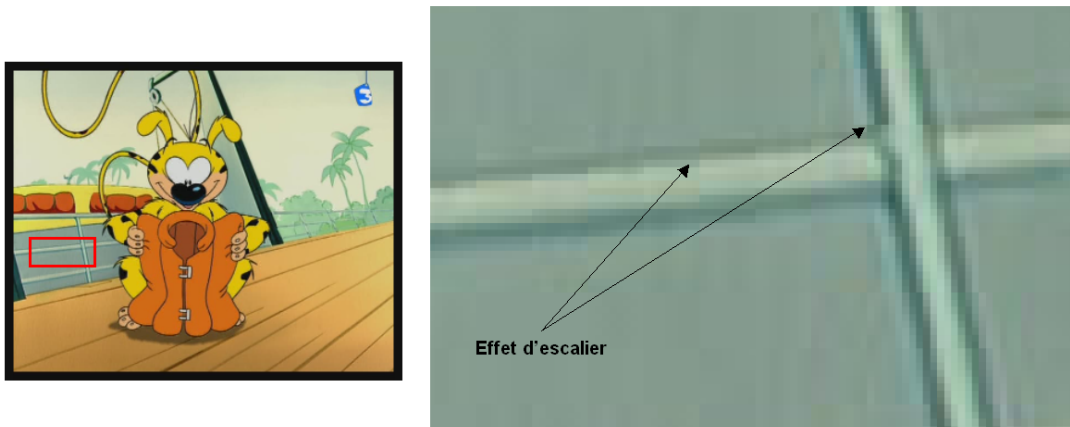


FIG. 1.17 – L'effet d'escalier.

La force de visibilité de l'effet d'escalier Plus le contour aura un angle faible par rapport à la normale, plus l'effet d'escalier sera visible. En effet, si le contour fait un angle de 45° avec la normale, il sera représenté par une combinaison de motifs n'appartenant pas à la première ligne (motifs horizontaux) ou la première colonne (motifs verticaux).

3.1.6 Le ringing

La cause La représentation d'un bloc est obtenue par la réunion de motifs DCT de telle sorte que la caractéristique de ces motifs soit mise en évidence, ou mise en défaut par la contribution des autres motifs. Dans le cas d'un contour d'un objet ou d'une personne, plusieurs motifs sont utilisés. Il est donc probable que la quantification tronque un ou plusieurs coefficients AC et introduise une erreur dans la reconstruction du contour.

La conséquence On voit alors apparaître le long des contours des ombres successives ou oscillations causées par la mauvaise combinaison des motifs basses fréquences utilisés pour le contour. C'est pourquoi ces oscillations sont généralement parallèles aux contours. Cet effet est appelé l'effet de ringing et la figure 1.18 illustre ce phénomène.



FIG. 1.18 – L'effet de ringing.

La force de visibilité de l'effet de ringing Plus le contour sera contrasté et proche d'une zone homogène, plus l'effet de ringing sera visible. Dans les autres cas, l'effet de masquage atténuera l'effet de ringing.

3.1.7 Les faux contours

La cause La quantification des blocs contenant peu d'informations fait tendre le contenu des blocs à une valeur homogène. Si on étend ce phénomène aux blocs adjacents d'une zone homogène, il est possible que ces valeurs homogènes varient d'un bloc à l'autre.

La conséquence La zone homogène non compressée qui pouvait contenir un dégradé de luminosité sera alors représentée par un dégradé par morceau. Un faux contour sera alors créé par la succession du phénomène de l'effet de bloc. La figure 1.19 représente cet artefact.

La force de visibilité des faux contours Le SVH a une sensibilité différente à la luminance. Plus le dégradé sera situé dans une zone où le SVH est sensible aux variations, plus le faux contour sera visible.

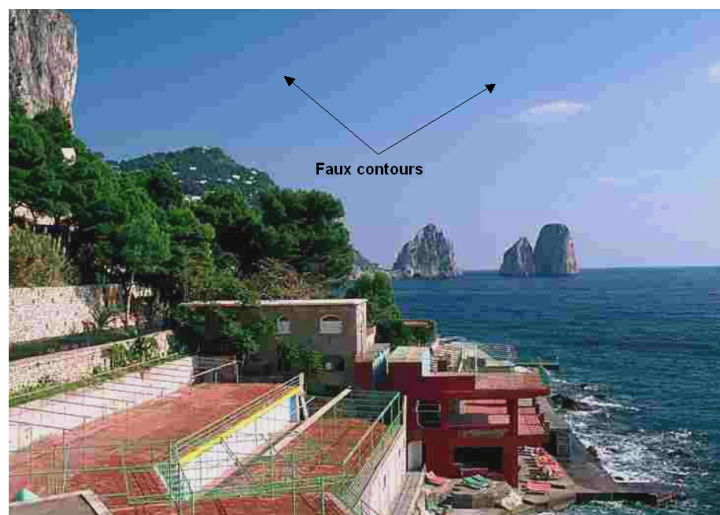


FIG. 1.19 – Les faux contours.

3.2 Les distorsions temporelles

3.2.1 L'effet de moustique

La cause La compensation de mouvement est obtenue par la combinaison de l'estimation de mouvement et du vecteur d'erreur de prédiction. Ces informations sont codées comme le contenu d'un bloc et soumis à la quantification. Une mauvaise estimation de mouvement ou une quantification trop importante du vecteur d'erreur pourront être responsables d'un codage différent d'une même information d'une image à l'autre.

La conséquence Cette même information codée différemment d'une image à l'autre causera une fluctuation autour des contours lorsqu'ils seront animés. Ce phénomène de battement a été comparé à des battements d'ailes de moustique ce qui explique sa dénomination par le terme d'effet de moustique. La figure 7.17 représente deux images consécutives et leur image différence calculée par la différence en valeur absolue des pixels de chaque image. Plus la différence est grande plus le pixel obtenu est foncé. Cela montre qu'autour des objets en mouvement, des fluctuations ont lieu et vont être perçues par le SVH comme des battements ou frémissements le long des contours.

La force de visibilité de l'effet de moustique L'effet de moustique sera d'autant plus visible sur des informations prédites déjà quantifiées. En effet, cette information sera difficile à prédire et le vecteur d'erreur sera conséquent et sera probablement quantifié.

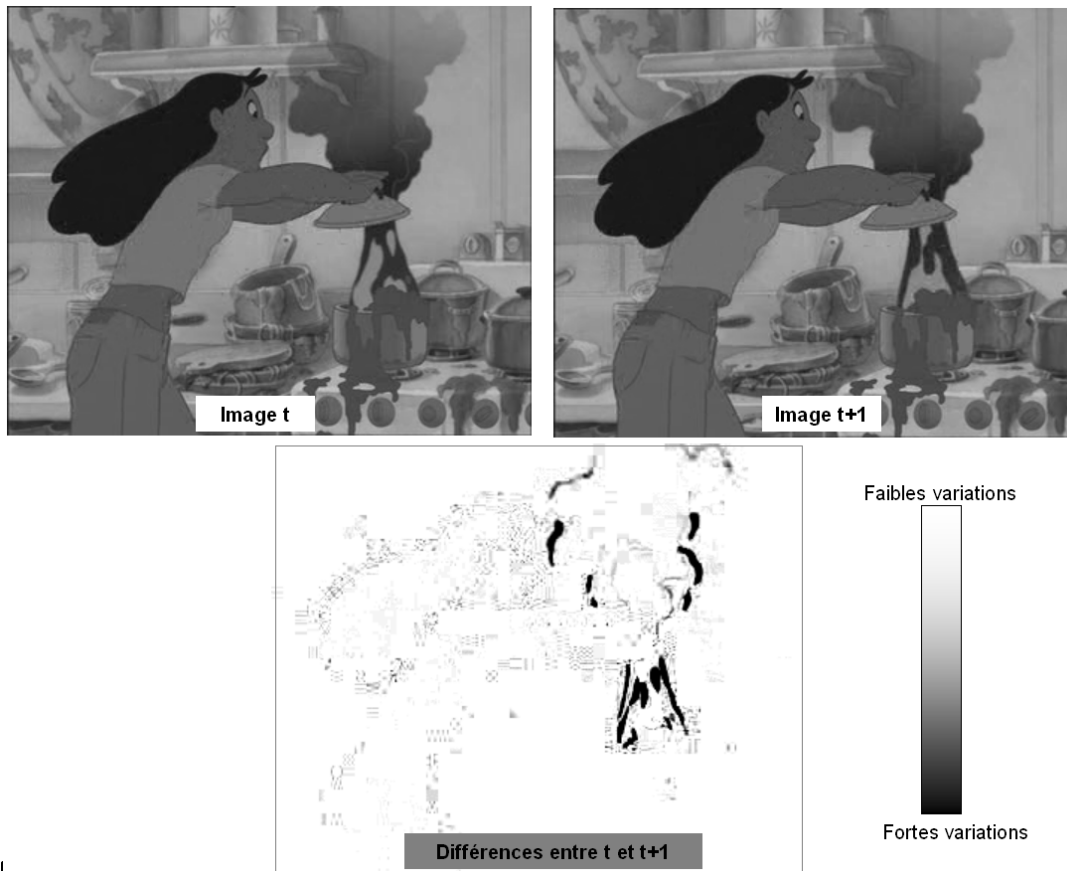


FIG. 1.20 – L'effet de moustique.

3.2.2 La perte d'informations

La cause Dans le cas d'une mauvaise estimation de mouvement, le vecteur d'erreur aura beaucoup d'information à transmettre. De plus, si ce vecteur est fortement quantifié, l'information sera non seulement transmise à la mauvaise position mais aussi avec un faux contenu.

La conséquence Lors de la reconstruction de l'image, l'information sera modifiée et pourra créer des disparités telles que la disparition d'information. Ce phénomène est représenté sur la figure 1.21 où le visage de la nageuse est coupé. On remarque d'ailleurs que cette coupure atteint tout un MB car la compensation de mouvement se fait par macrobloc.

3.2.3 Un mouvement saccadé

La cause La limite de la zone de recherche dans laquelle peut être estimé le vecteur de mouvement peut parfois être insuffisante.

La conséquence Dans ce cas, le mouvement ne sera pas représenté d'une image à l'autre et la fluidité du mouvement au cours du temps sera cassée provoquant le mouvement saccadé. Ce phénomène est difficilement représentable par illustration statique.

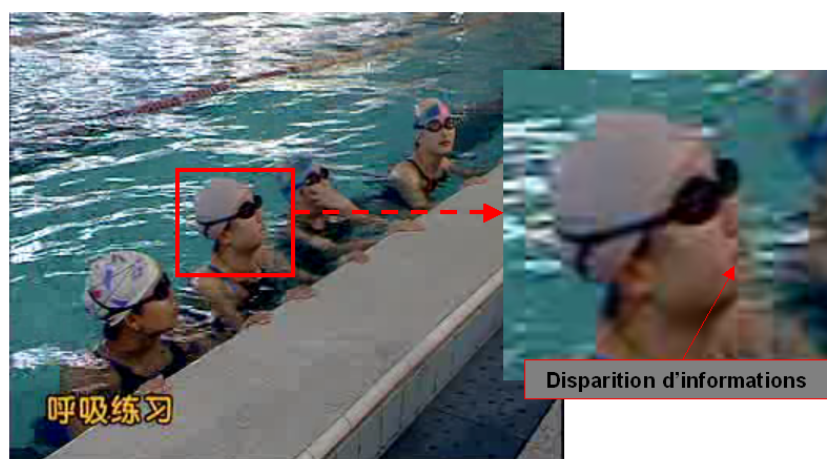


FIG. 1.21 – Exemple de disparités dues à la compensation de mouvement.

3.2.4 L'effet de scintillement ou flicker

La cause Le traitement image par image, associé au faible nombre de rafraîchissement par seconde d'un dispositif d'affichage, peut générer une discontinuité au niveau de la visualisation des images en temporel.

La conséquence Cela provoque un effet de scintillement, généralement appelé "flicker". Ce phénomène étant lié au dispositif d'affichage, n'est pas une conséquence directe de la compression. Cependant, il est important de le citer car il est souvent mentionné dans le domaine de la vidéo.

Le tableau 1.3 résume les différentes distorsions ou artefacts évoqués précédemment en précisant pour chacun si il est la conséquence d'un ou de plusieurs phénomènes utilisés pour la compression spatiale ou pour la compression temporelle, et si il apparaît au niveau d'un bloc, d'un macrobloc ou d'une zone de l'image.

TAB. 1.3 – Liste des artefacts causés par la compression spatiale et/ou temporelle, apparaissant au niveau du bloc et/ou du macrobloc

Nom de l'artefact	Spatiale	Temporelle	Bloc	Macrobloc
Effet de bloc	x		x	x
Effet de motif DCT	x		x	
Le flou	x	x	x	x
Les fausses couleur	x			x
L'effet d'escalier	x		x	
Le ringing	x		x	
Les faux contours	x			
Le mosquito		x		
La disparition d'information		x	x	
Les mouvements saccadés		x	x	
L'effet de scintillement		x		

Nous avons consacré ce chapitre à définir le rôle, les principes et les conséquences de la compression numérique. L'historique sur les normes de compression a permis de montrer que la majorité des techniques de compression s'appuyait sur deux mêmes lignes directrices qui sont l'étude de la redondance des informations au cœur d'une même image et l'étude de la redondance au sein de plusieurs images successives constituant une vidéo. De plus, nous avons situé la norme MPEG-2 en montrant que c'était actuellement la plus utilisée pour la télévision numérique.

Puis, nous avons expliqué les principes de compression utilisés dans la norme MPEG-2 et les précautions prises par les groupes de travail pour toujours tenir compte de la sensibilité du SVH et de sa capacité à percevoir un manque d'information dans l'image. En effet, la combinaison de méthodes de compression sans pertes et de compression avec pertes mais non perceptibles par le SVH permet d'obtenir des taux de compression non négligeables. Néanmoins, malgré ces précautions, lorsque la bande passante est trop restreinte, les pertes d'informations ne sont plus contrôlées et ne respectent plus ces caractéristiques liées à la perception humaine.

Nous avons alors listé les répercussions visuelles provoquées par ces pertes de données en mettant en avant les causes et les conséquences, dépendant non seulement des principes de compression, mais aussi de la bande passante et du contenu de la vidéo.

Si on se réfère à l'historique du début 1.1, la compression spatiale étant déjà présente dans la norme JPEG de compression d'image fixe, des artefacts tels que l'effet de bloc ou le ringing existent depuis la fin des années 80. Un large choix de méthodes a donc déjà été proposé pour éviter l'apparition de ces artefacts ou pour les corriger. C'est ce que nous proposons d'étudier dans le chapitre suivant.

Chapitre 2

Etat de l'art des corrections

Dans ce chapitre, nous proposons d'étudier les méthodes de la littérature qui nous permettront de réduire les artefacts de compression de type MPEG-2. Nous avons ciblé deux artefacts en priorité et nous commencerons par justifier ce choix. Puis, parmi les méthodes dédiées à ces artefacts, nous ciblerons une catégorie précise correspondant à notre contexte d'étude. Dans cette catégorie, nous décrirons les traitements proposés et nous sélectionnerons les plus adaptés à nos contraintes. Enfin, nous verrons que ces traitements peuvent être améliorés mais au risque de faire face à de nouveaux artefacts.

1 Le choix des artefacts et de la catégorie de correction

1.1 Le choix des artefacts

Les articles de la littérature proposent majoritairement des solutions destinées à supprimer l'effet de bloc et l'effet de ringing. Nos choix se portent également sur ces deux artefacts.

1.1.1 L'effet de bloc et l'effet de ringing : les artefacts les plus gênants

L'effet de bloc apparaît comme des frontières séparant des blocs d'information et l'effet de ringing apparaît comme du bruit le long des contours.

Même si ces artefacts sont la conséquence d'une perte d'information, ils sont perçus comme une information non naturelle ajoutée sur l'image, que le Système Humain a des difficultés à intégrer ou à corrélérer aux autres informations contenues dans l'image. Ces artefacts sont donc très gênants et prioritaires à corriger.

1.1.2 Les artefacts atténués par la correction de l'effet de bloc et de l'effet de ringing

Cette catégorie d'artefacts sera indirectement atténuée par la correction de l'effet de bloc et de l'effet de ringing. En effet, l'effet d'escalier est un artefact dont la visibilité est amplifiée par l'effet de bloc : il est d'autant plus visible que les frontières de blocs contenant chaque morceau de la droite à reproduire sont marquées. Les faux contours, eux, sont la conséquence de l'effet de bloc : ils sont créés par une succession de frontières de blocs dans des zones avec peu de détails. Ainsi, en supprimant l'effet de bloc, ces deux artefacts seront atténués.

L'effet de moustique, quant à lui, est causé par du bruit le long des contours qui varie d'une image à l'autre. Or, en corrigeant l'effet de ringing, le bruit le long des contours sera atténué. La conséquence est que l'effet de moustique sera indirectement atténué par la correction de l'effet de ringing.

1.1.3 Les artefacts moins prioritaires

Cette catégorie représente les artefacts qui sont moins prioritaires, soit car ils sont moins gênants que l'effet de bloc et l'effet de ringing, soit parce qu'ils nécessitent des corrections trop complexes.

Dans des proportions raisonnables, les artefacts tels que l'effet de motif, l'effet de flou, ou la disparition d'information sont des artefacts moins gênants. En effet, de part son vécu, le SVH est capable d'interpoler l'information manquante et n'est donc pas aussi sensible à ces artefacts qu'à l'effet de bloc ou l'effet de ringing.

En revanche, si l'effet de flou ou la disparition d'information est importante, le manque d'information sera gênant pour la compréhension de l'image. Dans ce cas, il sera difficile d'envisager leur correction avec une méthode simple et n'ayant aucun renseignement sur le contenu de l'image non compressée.

De même, pour l'artefact des mouvements saccadés, il pourrait éventuellement être détecté et corrigé en étudiant le contenu d'images successives. Des méthodes d'interpolation pourraient recréer le mouvement, mais la contrainte de temps réel serait difficilement envisageable.

Enfin, un artefact tel que **les fausses couleurs** apparaît comme une information fautive. Cette information est difficilement détectable par une méthode simple. En effet, l'étape de détection nécessite une corrélation entre les trois composantes [Catorina et al., 2007]. De plus, cet artefact apparaît lorsque la compression est très élevée. Dans ce cas, le SVH sera plus gêné par l'effet de bloc et l'effet de ringing (qui seront très prononcés) que par les fausses couleurs.

1.2 Le choix de la catégorie de correction

L'effet de bloc et l'effet de ringing. Ces deux artefacts peuvent être corrigés à plusieurs niveaux dans la chaîne de la diffusion numérique :

- Avant l'encodage par les méthodes de pré-traitements
- Pendant l'encodage par modification du principe de l'encodeur ou par traitement dans la boucle d'encodage
- Après le décodage par des méthodes dites post-traitements

1.3 Les méthodes de pré-traitements avant l'encodage

La compression des données favorise l'information à laquelle le SVH est le plus sensible. Dans le même état d'esprit, [Ozkan et al., 1993] et [Kleihorst et al., 1995] proposent des méthodes de pré-filtrage. Ces méthodes suppriment le bruit en utilisant des techniques de compensation de mouvement pour réduire l'information inutile. Cela permet de laisser plus de bits aux informations essentielles qui seront donc moins quantifiées.

Dans le même registre, l'idée proposée dans [Kutka et al., 1996] est d'accentuer le contraste de l'image avant l'encodage. Ainsi, comme les dégradations génèrent souvent une perte de contraste, ces pertes seront compensées par l'étape de pré-traitement. Ce procédé est illustré sur la figure 2.1.

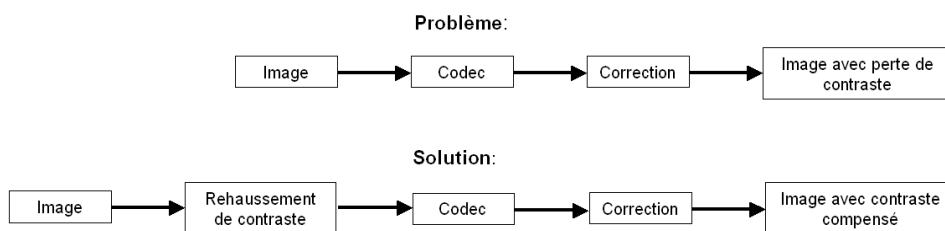


FIG. 2.1 – Principe de la méthode de Kutka et al : pré-traitement par rehaussement du contraste

Ces méthodes améliorent la qualité mais requièrent l'accès à l'image originale avant compression, ce qui est rarement le cas.

1.4 Les méthodes liées à l'encodage

1.4.1 Les modifications des techniques d'encodage

D'autres méthodes proposent une amélioration directement liée au système d'encodage pour éviter ou diminuer l'apparition des artefacts. Dans [Macq, 1992], la méthode proposée consiste à modifier les matrices de quantification en fonction du comportement du rapport signal sur bruit.

Dans [Macq et Shi, 1993] et [Nill, 1985], les auteurs proposent de combiner la transformée en cosinus discrète à un model basé sur le système visuel humain. Ces techniques, liées à notre perception, permettent de mieux quantifier les blocs selon leur contenu afin d'avoir le moins d'artefact possible.

Dans [Reeve et Lim, 1984], l'un des premiers articles proposés sur la réduction de l'effet de bloc, Reeves et Lim décrivent une solution qui modifie le schéma classique de la compression par blocs. Comme l'effet de bloc est causé par la quantification indépendante des blocs, l'idée est de corriger le problème à la source en divisant l'image en plusieurs blocs se chevauchant. Certains pixels se trouvent alors dans plusieurs blocs et lors de la reconstruction de l'image, un pixel codé un ou plusieurs fois prend la valeur moyenne des valeurs codées. Cette méthode élimine les effets de blocs sans détériorer l'image mais augmente le nombre de bits en sortie nécessaire au codage de l'information car il y a beaucoup plus de blocs à traiter.

Ces techniques de modification de l'encodage impliqueraient de devoir modifier la majorité des encodeurs et décodeurs du marché. Même si la solution de résoudre le problème à la base semble la plus logique, elle n'est pas réalisable dans notre contexte.

1.4.2 Les corrections dans la boucle d'encodage

Nous avons vu que la compression temporelle par compensation de mouvement se faisait à partir d'une image déjà compressée. L'intérêt d'introduire une correction dans la boucle d'encodage est d'améliorer la qualité des images de référence utilisées pour l'estimation de mouvement. A ce niveau, plusieurs types de méthodes ont été proposées. Dans [Pang et Tan, 1997], les auteurs proposent une faible correction au niveau des frontières donnant un meilleur PSNR mais insuffisant pour corriger efficacement les frontières de blocs.

Lee et Park dans [Lee et al., 1998] et [Lee et Park, 2001] proposent une méthode plus élaborée de filtrage. Leur méthode permet de corriger l'effet de bloc et l'effet de ringing en se basant sur les coefficients DCT, le type de macrobloc et le paramètre de quantification.

Dans le même état d'esprit, dans l'encodeur MPEG-4/H264, les auteurs [List et al., 2003] se basent sur les coefficients AC et DC des blocs, sur les paramètres du décodeur et proposent une correction efficace et robuste mais uniquement une correction de l'effet de bloc.

Ces méthodes rendent robustes les corrections. En effet, ces paramètres renseignent sur la force de compression et donc donnent une idée de la force de correction à appliquer ce qui explique l'intérêt des recherches portées sur ces méthodes jusqu'à aujourd'hui [Beermann et Wien, 2004], [Beermann et Wien, 2005], [Sun et al., 2006].

Tout comme pour les méthodes de pré-traitements, notre contexte ne nous permet pas d'agir à l'intérieur du décodeur ce qui rend les méthodes de corrections dans la boucle d'encodage inexploitable.

1.5 Les méthodes post-traitements après le décodage de l'image

Ces méthodes agissent directement sur les images décodées et n'ont accès qu'aux valeurs des pixels et parfois aux paramètres du décodeur. Ces traitements sont plus conséquents en temps de calcul qu'un traitement dans la boucle mais pourront toujours s'adapter sur les images quels

que soient les paramètres de l'encodage. Cela augmente considérablement le domaine d'application de ce type de méthodes, et c'est pourquoi elles ont été majoritairement étudiées. On trouve d'ailleurs dans la littérature, des articles [Shen et Kuo, 1997] et [Nosratinia, 2001] entièrement consacrés à l'état de l'art de ces corrections.

Parmi ces corrections post-traitement, il existe deux techniques distinctes : la technique de restauration d'images et la technique d'amélioration d'images.

1.5.1 Les méthodes de restauration

L'objectif final de la restauration d'images est de modifier une image pour obtenir l'image la plus proche possible de ce qu'elle devrait être en l'absence de dégradations. La figure 2.2 schématise ce principe de restauration. A partir de l'image dégradée y , des connaissances sur l'image non dégradée x et/ou sur sa compression, l'image restaurée x' sera reconstruite. Pour ce faire, la méthode modélise la dégradation et applique la dégradation inverse à l'image dégradée pour retrouver l'image originale. L'image obtenue après restauration traduit de la façon la plus fidèle possible la scène qu'elle représente, même si cela conduit à une image de faible qualité.

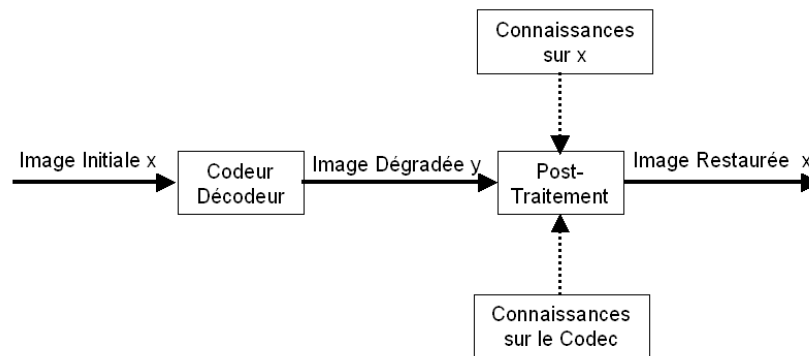


FIG. 2.2 – Principe général des méthodes de restauration.

Pour modéliser la dégradation et appliquer le phénomène inverse, ces méthodes se basent sur la minimisation d'une métrique ([Wu et Gersho, 1992], [Nakajima et al., 1994], [Gopinath et al., 1994], [Hong et al., 1998]), la maximisation d'une probabilité ([Ozcelik et al., 1995], [O' Rourke et Stevenson, 1995] et [Yang et al., 2001]) ou sur un ensemble de contraintes ([Zakhor, 1992], [Yang et al., 1993], [Su et Mersereau, 1995] et [Paek et al., 1998]).

Plus récemment [Bourdon et al., 2003], une technique utilisant les équations aux dérivées partielles ou EDP a été développée. Elle permet de restaurer les images compressées en JPEG2000. Cette technique corrige les effets de distorsion autour des discontinuités d'une image pour des bas débits, tout en préservant ses contours. Ce type de méthode peut donc être envisagé pour la correction de l'effet de ringing.

Ces techniques sont réputées pour leur efficacité mais nécessitent un temps de calcul trop important causé par des itérations successives. De plus, n'ayant accès ni aux connaissances sur l'image ni sur l'encodeur, nous ne pourrions pas les exploiter.

1.5.2 Les méthodes d'amélioration

L'amélioration d'images est un domaine qui regroupe tout un ensemble d'opérations réalisées sur les images avec comme objectif général d'obtenir une image de meilleure qualité visuelle. Elle ne vise pas la fidélité de la scène mais la bonne qualité visuelle de l'image finale même si celle-ci est différente de l'originale. Ce type de méthode est donc exploitable dans notre contexte, c'est pourquoi, nous allons les détailler dans la section suivante.

La figure 7.19 résume les points de cette première section. Après avoir choisi de privilégier l'effet de bloc et l'effet de ringing, nous avons sélectionné la méthode d'amélioration d'image en post-traitement pour pouvoir réduire les artefacts de compression sur l'image décodée.

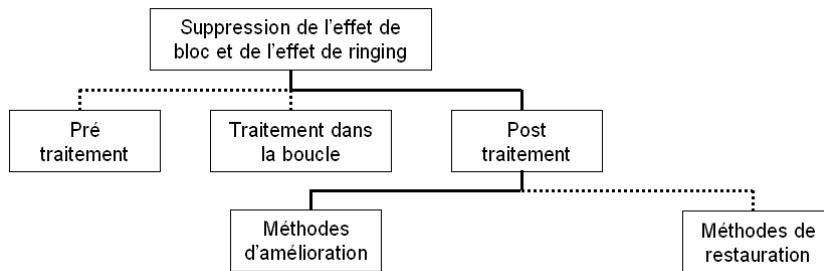


FIG. 2.3 – Cheminement suivi pour choisir la catégorie de méthode de correction.

2 Etat de l'art des méthodes d'amélioration d'images

Nous allons étudier les méthodes d'améliorations d'images proposées dans la littérature pour supprimer l'effet de bloc et l'effet de ringing. Le but de ces corrections est de supprimer l'information non naturelle ajoutée par ces deux artefacts. La suppression de cette information est donc principalement réalisée par des techniques de filtrage. Néanmoins, d'autres principes de correction ont été proposés et nous commencerons par une description de ces méthodes inhabituelles. Puis, nous reviendrons aux techniques de filtrage et nous verrons pourquoi et comment ces méthodes ont dû évoluer. Nous verrons que ces méthodes, d'abord dédiées à la correction de l'effet de bloc, ont également été utilisées pour la correction de l'effet de ringing.

2.1 Des méthodes inhabituelles

Que ce soit pour l'effet de bloc ou pour l'effet de ringing, la majorité des problèmes provient de la perte des coefficients hautes fréquences lors de la quantification. Dans [Niss, 1998], la technique consiste à retrouver ces coefficients hautes fréquences AC. En utilisant le coefficient DC d'un bloc donné et les coefficients DC des huit autres blocs voisins, Niss arrive à prédire par interpolation polynomiale les 5 derniers coefficients AC du bloc considéré. Cette méthode fait ses preuves dans les zones monotones mais donnent de moins bons résultats lorsque les zones contiennent des contours car elles possèdent plus de coefficients AC. Néanmoins, si on tient compte de l'effet de masquage, c'est dans les zones avec peu de coefficients AC que les dégradations sont les plus visibles. La méthode nous semble donc très bien adaptée pour corriger les artefacts gênants pour le SVH. Par contre, le re-calcul des coefficients AC ne correspond pas à notre contrainte de faible coût.

Dans un autre registre proche de la solution de pré-traitement proposée par Reeves et Lim [Reeve et Lim, 1984] sur le chevauchement des blocs, Nosraturia et al [Nosratinia, 2001] proposent de réappliquer le schéma de compression JPEG à l'image compressée. L'application étant réservée aux images JPEG, ils récupèrent les informations de compression se trouvant dans l'en-tête du fichier et appliquent cette même compression sur l'image décalée dans huit directions différentes. Enfin, en moyennant toutes ces images, le résultat obtenu est l'image compressée dont l'effet de bloc est largement réduit.

Cette idée, exploitée ici pour les images JPEG est réutilisée par Samadani et al [Samadani et al., 2004] qui proposent une amélioration de la correction de l'effet de ringing en modifiant la compression pour les pixels proches des contours. La réduction de l'effet de ringing est notable.

Ces approches sont originales car elles n'utilisent pas le schéma classique de correction que nous allons décrire dans la suite. Néanmoins, ces méthodes sont difficilement exploitables.

2.2 Des techniques de filtrage linéaires et non linéaires

Le principe du filtrage consiste soit à supprimer de l'information gênante soit à faire ressortir de l'information importante. Cela correspond à notre besoin pour corriger l'effet de bloc et l'effet de ringing.

Le filtrage Filtrer spatialement une image consiste à modifier la valeur d'un pixel proportionnellement aux variations d'intensité lumineuse de ses voisins. Ces filtres sont utilisés pour augmenter certains contrastes, détecter des fronts ou des contours, ou au contraire réduire le bruit par lissage. Si $I(i,j)$ représente l'intensité du pixel I de coordonnées (i,j) dans l'image, les pixels entourant $I(i,j)$ peuvent être indexés comme indiqué sur la figure 2.4 (dans le cas d'un voisinage 3*3).

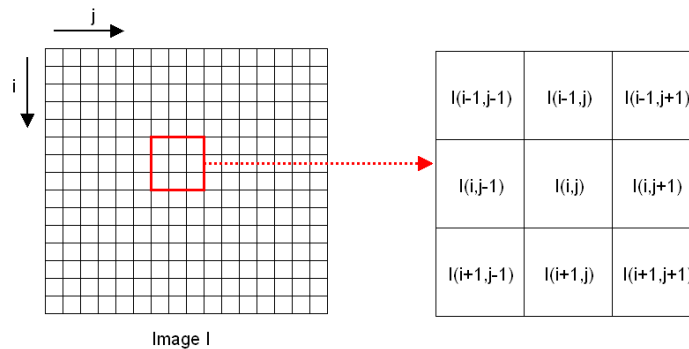


FIG. 2.4 – Schématisation des pixels d'une image I et des coordonnées de ces pixels dans un voisinage 3*3

Un filtre linéaire affectera à $I(i,j)$ une valeur qui est une combinaison linéaire des pixels voisins suivant un voisinage donné. Un filtre non linéaire affectera à $I(i,j)$ une valeur qui n'est pas une combinaison linéaire des pixels voisins suivant un voisinage donné.

2.2.1 Les filtres linéaires ou de convolution

Une convolution est une fonction mathématique qui remplace chaque pixel par une somme pondérée de ses voisins. La matrice définissant la taille du voisinage contient les poids respectifs de chaque pixel voisin. Cette matrice est aussi appelée masque ou noyau de convolution. Filtrer un signal 2D (une image) consiste à convoluer sa fonction d'intensité I avec une fonction h appelée réponse impulsionnelle du filtre :

$$I'(i,j) = h(i,j) * I(i,j). \quad (2.1)$$

Pour supprimer les hautes fréquences représentées par les frontières de blocs, l'idée est d'utiliser un filtre passe-bas pour ne laisser passer que les basses fréquences. Par exemple, en utilisant le filtre moyenneur qui moyenne toutes les valeurs autour du pixel central, les hautes fréquences sont atténuées par le poids des basses fréquences.

La première correction utilisée par Reeves et Lim [Reeve et Lim, 1984] en 1984 proposait l'application d'un filtre passe-bas de type gaussien. Ce filtre a un effet assez similaire au filtre moyenne, mais la moyenne est pondérée, en ce sens où les pixels près du centre ont un poids plus important que ceux qui sont situés plus loin. La fonction gaussienne utilisée dans le traitement d'image est une fonction à deux dimensions. Reeves et Lim proposent d'utiliser ce filtre Gaussien avec les coefficients énoncés sur la formule ci-après.

$$h(i, j) = \exp[-0.5(i^2 + j^2)] \quad \text{avec } i = \{1, 0, -1\}, j = \{1, 0, -1\} \quad (2.2)$$

soit

$$h(i, j) = \begin{bmatrix} 0.0751 & 0.1239 & 0.0751 \\ 0.1239 & 0.2042 & 0.1239 \\ 0.0751 & 0.1239 & 0.0751 \end{bmatrix} \quad (2.3)$$

Malheureusement, ce procédé lisse effectivement les frontières mais lisse également les contours des objets de l'image se trouvant au niveau des frontières des blocs.

En 1994, Jarske et al [Jarske et al., 1994] proposent d'examiner l'efficacité des différents filtres à supprimer les frontières de blocs tout en gardant le contraste dans l'image. Ils concluent qu'un filtre gaussien suivi d'un rehaussement de contours semble être le plus efficace. En effet, si un filtre gaussien a tendance à trop lisser les contours, un filtre rehausseur de contours aura l'effet inverse. L'étape de rehaussement de contours est effectuée à l'aide d'un filtre passe-haut. A l'inverse d'un filtre passe-bas, le filtre de rehaussement de contours localise les hautes fréquences et les accentue. La formule suivante donne le noyau du filtre utilisé par les auteurs :

$$h(i, j) = \begin{bmatrix} 0 & -\alpha & 0 \\ -\alpha & 1 + 4\alpha & -\alpha \\ 0 & -\alpha & 0 \end{bmatrix} \quad (2.4)$$

Cette technique de filtrage et de rehaussement de contours est un très bon compromis. Néanmoins, si les hautes fréquences ne correspondant pas aux frontières de blocs mais à de l'information dans l'image sont trop filtrées par un passe-bas, ces informations ne pourront pas être rehaussées. Inversement, si les frontières de blocs ne sont pas assez lissées car le filtre passe-bas utilisé est trop faible, elles seront rehaussées par le filtre passe-haut.

Sur la figure 2.5, une image compressée A est filtrée par l'intermédiaire d'un filtre gaussien. Par rapport à A les frontières de blocs sont atténuées mais l'image est floutée. Les images C et D représentent l'image B rehaussée avec le filtre présenté pour différentes valeurs de alpha. On note sur C que l'image est moins floutée et que les frontières sont moins visibles que sur A. Sur l'image D, il est clair que l'image est très nette mais toutes les frontières de blocs et le ringing qui n'avaient pas totalement été atténuées se retrouvent totalement rehaussées. La difficulté est de trouver le meilleur compromis sachant qu'il sera variable selon le contenu de l'image.



FIG. 2.5 – Exemple de correction de type filtrage et rehaussement. α est le coefficient du filtre de Jarske et al dont le noyau est représenté à l'équation (2.6).

2.2.2 Les filtres non linéaires

En constatant que les frontières de blocs sont liées à l'information dans l'image, Ramamurthy et al [Ramamurthy et Gersho, 1986] ont proposé en 1986 de remplacer les filtres linéaires par des filtres non linéaires pour pouvoir mieux tenir compte des contours et moins les lisser. En 1993, Hsu et al [Hsu et Chen, 1993] proposent un filtre médian non-linéaire à deux dimensions. En recherchant la valeur médiane des coefficients d'un noyau, ce filtre peut atténuer les pixels isolés d'une valeur très différente de leur entourage. Ce filtre médian non-linéaire a la caractéristique de bien préserver les contours mais reste toujours confronté au problème de différenciation entre les contours et les frontières de blocs.

2.2.3 Les inconvénients de ces techniques de filtrage

L'effet de flou Le rôle de ces différents filtres est de supprimer l'information non naturelle pour le SVH. L'inconvénient de ces filtres est qu'ils ne sont pas localisés et qu'ils appliquent le même principe sur toute l'image. La conséquence est la perte de détails sur tout l'image créant l'information de flou.

L'accentuation de l'effet de bloc Pour éviter ce phénomène de flou, certains filtres sont plus appropriés pour ne pas dégrader les contours. Dans ce cas, les frontières seront détectées

comme des contours et ne bénéficieront pas d'un filtrage plus agressif. La conséquence est que le contenu des blocs va être plus filtré que la frontière. Ainsi, l'effet de masquage contenu dans les blocs sera supprimé rendant ainsi plus gênante la perception de l'effet de bloc.

Pour ces raisons, ces techniques de filtrage ont rapidement dû être localisées pour ne pas dégrader l'image. C'est pourquoi, des techniques de classification précédant le filtrage ont été proposées.

2.3 Les corrections par classification et filtrage

Avec l'apparition des différentes normes de compression d'images fixes et de vidéos, beaucoup de méthodes ont été proposées. On distingue dans l'étape de classification la technique de classification par régions et celle de classification par zones.

2.3.1 La classification par région

En 1996, Meier et al [Meier et al., 1996] proposent une classification par segmentation de l'image. La segmentation est effectuée par la technique de croissance de région dans le but d'isoler des contours. Puis, les pixels situés à l'intérieur des zones délimitées par les contours sont filtrés par un filtre gaussien 3×3 . Pour éviter de dégrader un contour, tous les pixels du noyau du filtre ne doivent pas appartenir au contour comme cela est représenté sur la figure 2.6.

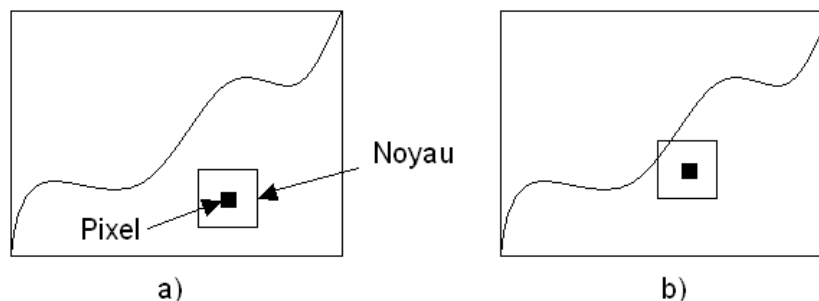


FIG. 2.6 – Les limites du filtrage dans une région définies par Meier et al : le filtrage du pixel considéré ne sera effectué que dans le cas a).

En 1998, Lee et al [Lee et al., 1998] font également une classification par zones. L'application est dédiée aux images JPEG. La classification est effectuée par le calcul d'un gradient sur toute l'image. Puis, les contours sont définis à partir d'un seuil global et d'un seuil local. Le seuil global dépend du coefficient de quantification utilisé pour l'image et le seuil local est déterminé par le calcul de la moyenne et de l'écart type des pixels appartenant à un bloc 8×8 . Puis, un filtre adaptatif à une dimension est appliqué dans les zones avec des contours tandis qu'un filtre à 2 dimensions est réservé aux zones homogènes.

Ces deux types de méthodes présentent l'inconvénient de devoir parcourir une première fois l'image pour l'analyser, puis une seconde fois pour effectuer les traitements. Cela oblige à garder en mémoire des blocs déjà étudiés et ne correspond pas à notre contrainte de faible coût.

Un autre type de classification par zones a été proposé par Xiong et al [Xiong et al., 1997] en utilisant des transformées en ondelettes. En effet, la transformée en ondelettes a pour particularité de décomposer le signal en une image avec des basses fréquences accompagnée de détails de plus en plus fins. Cela permet une classification progressive des zones homogènes et des zones contenant de l'information. Les choix de filtres sont ensuite déterminés par le contenu des régions. Malgré son efficacité, cette étude nécessite le calcul de la transformée en ondelette difficilement réalisable à faible coût.

2.3.2 La classification par bloc

Pour pouvoir exploiter les méthodes de classification et filtrage dans un décodeur, les méthodes de classification par bloc se sont développées. En effet, le parcours de l'image bloc par bloc est la technique utilisée pour répondre à la contrainte de temps réel : pour un bloc de 8 pixels par 8 pixels, 8 lignes seront lues mais seulement 64 coefficients seront gardés dans une mémoire.

Dans [Kundu, 1995], Kundu propose de classer les pixels des blocs en 3 catégories : les pixels présents dans des régions homogènes, les pixels étant proches des contours de l'image et les pixels se situant dans les régions texturées (figure 2.7 a)). C'est le type de classification le plus fréquemment utilisé car la classification est faite en même temps que le parcours des blocs. Cela permet également d'appliquer directement la correction en fonction du contenu du bloc. Sur la figure 2.7 b), un exemple de la longueur d'un filtre symétrique à une dimension est représenté pour corriger l'effet de bloc dans les trois types de régions. Ce type de filtre à une dimension est souvent utilisé pour les corrections temps réel qui appliquent leur filtrage ligne par ligne.

Remarque : le fait que le filtre soit symétrique limite le nombre total de filtres possibles à appliquer et réduit ainsi le coût de l'implémentation. L'inconvénient de l'aspect symétrique est que la correction étant appliquée sur le même nombre de pixels à gauche et à droite de la frontière, l'image corrigée peut présenter un aspect symétrique non naturel dans l'image et donc gênant pour le SVH.

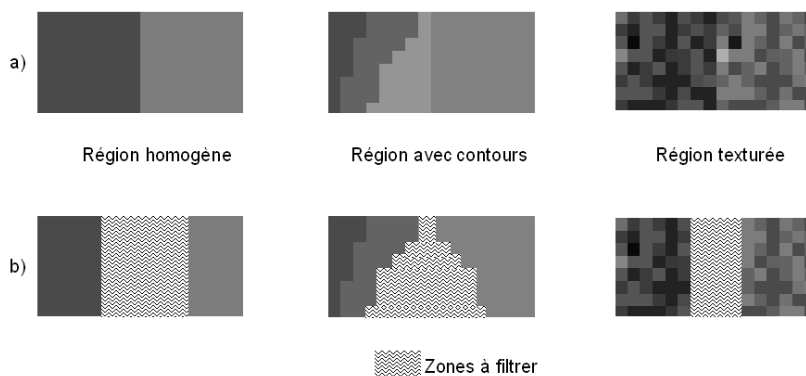


FIG. 2.7 – a) Représentation de différents contenus dans une paire de blocs. b) Représentation des zones de filtrage qui s'adaptent au contenu des blocs.

C'est ce type de méthode que nous souhaitons exploiter pour réduire les artefacts de compression de type MPEG-2. Néanmoins, comme la méthode de Kundu, nous allons voir que la classification est souvent obtenue au terme de calculs statistiques, et de seuils dépendant de caractéristiques plus globales de l'image.

Dans [Kuo et Hsieh, 1995], les auteurs proposent de définir la visibilité V d'un bloc en analysant la variation d'information autour de la frontière de la façon suivante :

$$V_{ij} = \sum_{k=1}^n |f_{ni+1,nj+k} - f_{nj,nj+k}| + |f_{n(i+1),nj+k} - f_{n(i+1),nj+k}| + \quad (2.5)$$

$$|f_{ni+k,nj+1} - f_{ni+k,nj}| + |f_{ni+k,n(j+1)+1} - f_{ni+k,n(j+1)}|$$

où V_{ij} est la visibilité du $(i, j)^{me}$ bloc de taille $n \times n$ et f_{pq} représente la valeur du pixel de l'image f à la position (p, q) . Puis, ils calculent la moyenne E et l'écart type σ de l'ensemble des V_{ij} obtenus pour chaque bloc. Enfin, si

$$V_{ij} > V_t = E(V) + \sigma(V) \quad (2.6)$$

alors le bloc est considéré comme visible et une correction peut être envisagée.

Cette méthode nécessite également de parcourir l'image entière pour obtenir la valeur seuil V_t et ne sera pas exploitable dans notre application. Néanmoins, leur méthode décrit un moyen pour éviter le filtrage dans un bloc contenant de l'information.

Une fois que ces frontières visibles sont détectées, il faut écarter les blocs contenant les contours pour ne pas les dégrader lors de l'application du filtre. Kuo et Hsieh [Kuo et Hsieh, 1995] proposent l'application de la combinaison des filtres de Robinson de type passe-haut.

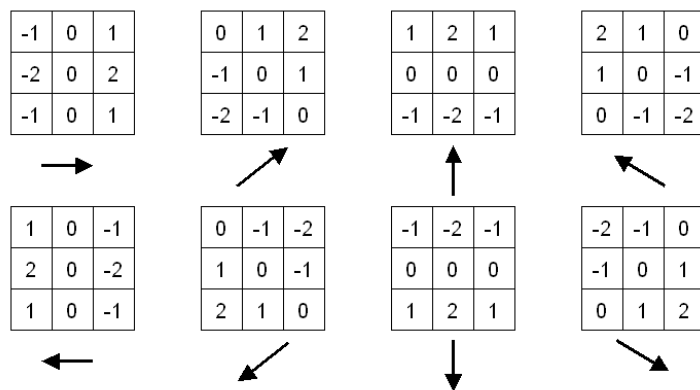


FIG. 2.8 – Les filtres de Robinson

Ces filtres incluent les filtres de Sobel capables de détecter les contours verticaux et horizontaux et les adaptent aux directions intermédiaires. Ces filtres passe-haut font donc ressortir les contours de l'image avec une forte intensité et les zones homogènes avec une faible intensité. L'inconvénient majeur est de trouver, à partir de quelle intensité appelée seuil, on définit que l'information fait partie d'un contour. Pour résoudre ce problème, les auteurs utilisent la construction d'un arbre

de contours basé sur la méthode de Lin et al [Lin et al., 1992]. Puis, pour les frontières n'étant pas détectées comme des contours, ils appliquent un filtre dont les coefficients dépendent de leur position de part et d'autre de la frontière. La figure 2.9 montre les coefficients du filtre utilisé selon la position sur la frontière.

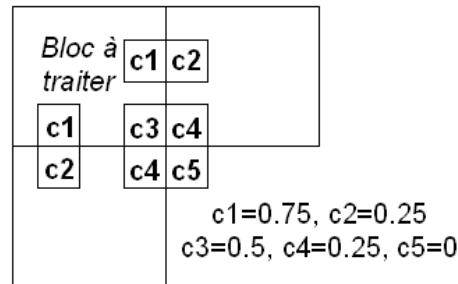


FIG. 2.9 – Exemple de filtrage passe-bas au niveau des frontières de blocs : méthodes de Lin et al

En 1997, Hu et al [Hu et al., 1997] utilisent le principe de classification pour réduire non seulement l'effet de bloc mais aussi l'effet de ringing. La différence par rapport à la méthode étudiée précédemment est que la classification se fait dans le domaine fréquentiel en utilisant les coefficients DCT. Cette technique est plus robuste mais plus coûteuse car il faut recalculer les coefficients DCT. Ensuite, ils effectuent un filtrage dans les zones homogènes c'est-à-dire dans les zones où l'effet de bloc se voit le plus et pour lesquelles les risques de dégradations par correction sont faibles.

Pour l'effet de ringing, comme il apparaît dans des zones proches des contours, ils utilisent un opérateur de Sobel pour les détecter et appliquent ensuite un filtre passe-bas le long de ces contours. L'utilisation de ces coefficients donnant une solution plus robuste, d'autres méthodes utilisant ce même principe ont été proposées. Dans [Luo et Ward, 2003], les auteurs ciblent une méthode qui préserve les contours et utilisent les coefficients DCT pour étudier la corrélation entre les intensités à la frontière des deux blocs pour distinguer les zones homogènes des zones avec des contours.

L'étape de classification peut être effectuée dans le domaine spatial ou dans le domaine fréquentiel en utilisant les coefficients DCT. Dans notre contexte, l'utilisation des coefficients DCT nécessite leur re-calcul, ce qui représente une étape ne correspondant pas à la contrainte de faible coût.

Nous venons d'étudier les différentes catégories de méthodes utilisées pour la correction de l'effet de bloc et de l'effet de ringing. A partir des méthodes d'améliorations d'images imposées par notre contexte d'étude, les méthodes de classification et filtrage dans le domaine spatial sont les plus adaptées à nos contraintes d'implémentation. Ce choix parmi les autres méthodes est résumé sur la figure 2.10.

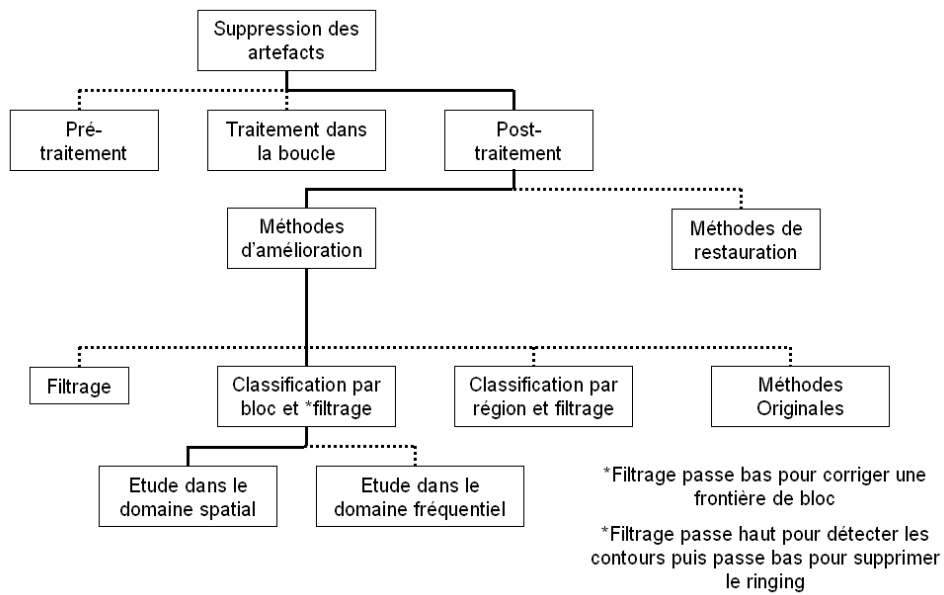


FIG. 2.10 – Cheminement suivi pour choisir la méthode de correction pour la réduction de l'effet de bloc et de l'effet de ringing.

Pour pouvoir améliorer la réduction des artefacts de compression MPEG-2, nous devons pouvoir adapter ces méthodes pour les rendre compatibles à nos contraintes d'implémentation. Nous allons détailler les difficultés rencontrées et les conséquences qu'elles entraînent.

3 Les limitations des méthodes de classification et filtrage

Dans notre contexte, la correction de l'effet de bloc et de l'effet de ringing par la technique de classification et filtrage peut être la cause de 3 catégories de problèmes. Nous allons lister ces problèmes et étudier si il existe des solutions dans la littérature que nous pourrions exploiter.

3.1 L'apparition des fausses frontières

L'apparition des fausses frontières vient de la correction de l'effet de bloc et de la contrainte de temps réel pour l'implémentation.

La cause La technique de classification et filtrage se divise en deux étapes : une étape de détection de frontière qui consiste à analyser les valeurs des pixels de chaque coté de la frontière, et, une étape de correction qui choisit en fonction de cette analyse le filtre le mieux approprié à la correction.

Pour appliquer cette méthode, une correction en temps réel effectue les étapes suivantes :

- Etape1 : Sélection de deux blocs Blk_n et Blk_{n+1} adjacents à la frontière B_n
- Etape2 : Analyse du comportement des pixels au niveau de la frontière B_n et à l'intérieur des blocs Blk_n et Blk_{n+1}
- Etape3 : Correction et modification des pixels des blocs Blk_n et Blk_{n+1} pour supprimer la frontière B_n

Puis l'étape1 recommence par la sélection des deux blocs Blk_{n+1} et Blk_{n+2} adjacents à la frontière B_{n+1} et ainsi de suite.

Dans ce type de méthode, il faut faire attention au problème de chevauchement entre les pixels à corriger et les pixels qui viennent d'être corrigés. En effet, la figure 2.11 illustre ce phénomène : par exemple, pour analyser la frontière B_{n+1} , nous avons besoin d'analyser les pixels du bloc Blk_{n+1} qui ont pu être modifiés lors de la correction de la frontière B_n . Pour éviter le problème de chevauchement entre les pixels corrigés et les pixels d'entrée, le nombre de pixels utilisé pour la détection et la correction des frontières est limité à la moitié de la taille de chaque bloc autour de la frontière. Ainsi, pour corriger la frontière B_{n+1} , l'analyse se fera sur la moitié droite du bloc Blk_{n+1} non modifiée par la correction de la frontière B_n .

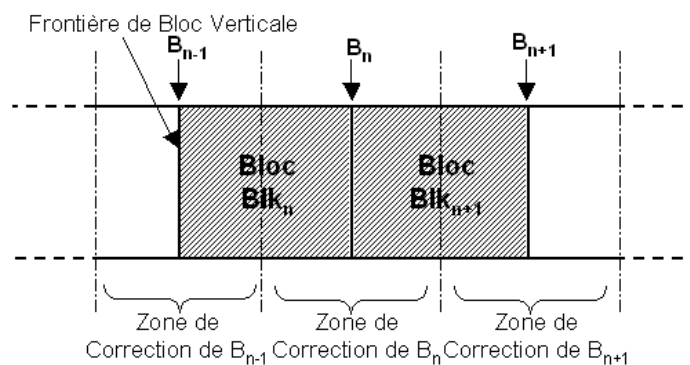


FIG. 2.11 – Principe traditionnel appliqué pour la correction de l'effet de bloc.

La conséquence Après une compression par bloc, les zones avec peu de détails sont représentées par une succession de blocs homogènes. Les corrections couramment utilisées consistent à filtrer la frontière sur la zone autorisée que nous venons de définir. La conséquence est l'apparition de frontières fantômes G_{n_G} et G_{n_D} à gauche et à droite de la frontière initiale B_n aux positions qui délimitent le début et la fin de la zone de filtrage. Ce phénomène est représenté sur la figure 2.7. Même si ces frontières fantômes sont moins perceptibles que la frontière initiale, elles rappellent tout de même le phénomène de structure artificielle gênante pour le SVH.

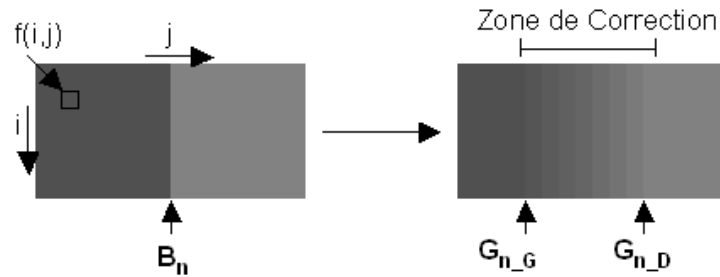


FIG. 2.12 – Le phénomène de création de fausses frontières ou frontières fantômes notées G_{n_G} et G_{n_D} .

Cas des blocs étendus De plus, lorsque certaines zones avec peu de détails sont fortement compressées, les valeurs des pixels de zones voisines convergent vers une même valeur. Ce phénomène provoque le regroupement de blocs entre eux et génère l'apparition de *blocs étendus* sans frontières intermédiaires. La figure 2.13 montre ce phénomène de regroupement pour les forts taux de compression : le nombre de frontières est moindre mais elles sont beaucoup plus perceptibles.

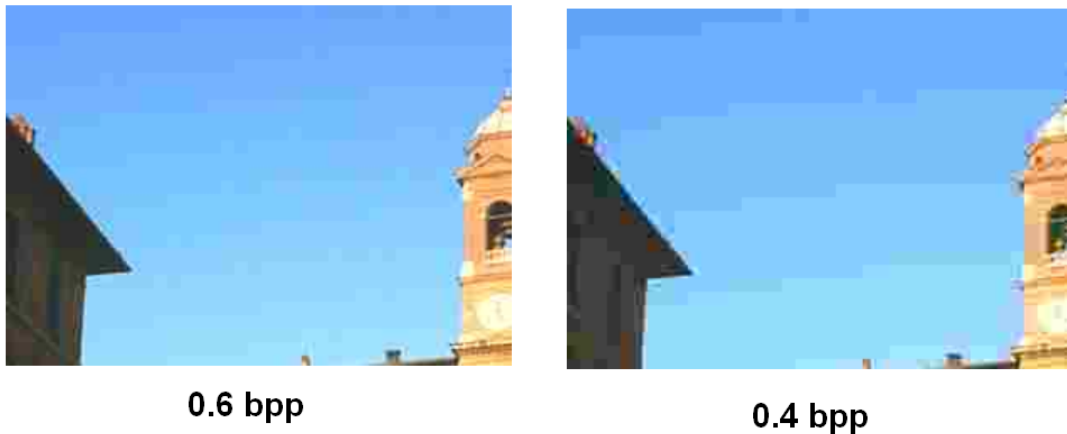


FIG. 2.13 – Le phénomène des blocs étendus : les frontières de blocs sont moins nombreuses sur l'image de droite qui est pourtant plus compressée (0.4 bits par pixels) que l'image de gauche (0.6 bits par pixels)

Dans ce cas, il est facile d'imaginer que les frontières fantômes résultantes d'une correction de l'effet de bloc seront très gênantes car très artificielles pour le SVH. La figure 2.14 montre l'exemple d'une correction temps réel que nous avons implémenté. Cette correction est effectuée

sur les blocs homogènes en appliquant un filtrage de type moyeneur limité à 4 pixels à gauche et 4 pixels à droite de la frontière. On voit que les frontières sont encore bien présentes car une structure artificielle et géométrique reste largement discernable par le SVH.

Remarque : nous avons volontairement représenté ces images en luminance pour ne pas inclure d'autres artefacts et cibler uniquement les frontières fantômes.

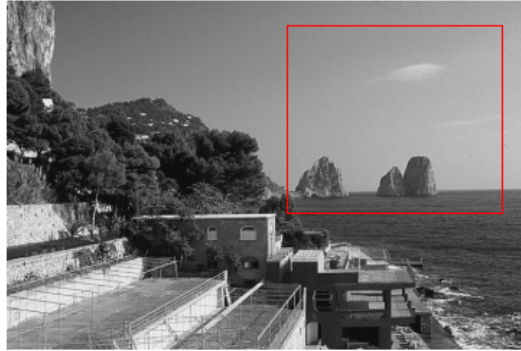


FIG. 2.14 – De gauche à droite : image non compressée, image compressée , image corrigée par filtre moyeneur sur 4 pixels à gauche et à droite de la frontière (les frontières fantômes sont visibles).

Les méthodes utilisées pour y remédier Des méthodes ont été proposées pour adapter les corrections des images fortement compressées dans le cas où un simple filtrage n'est pas suffisant. Dans [Hu et al., 1997], la méthode est la suivante : les blocs en question sont représentés par leur coefficient DC. La probabilité d'avoir un coefficient DC différent pour chaque bloc augmente la probabilité que les frontières de blocs apparaissent entre les blocs voisins. Dans ce cas, une correction en deux étapes peut supprimer ces frontières. La première étape consiste à détecter 4 blocs voisins n'ayant plus qu'un coefficient DC et à les regrouper en un macrobloc à l'intérieur duquel aura lieu le traitement. Ce traitement est un principe récursif qui consiste à appliquer à chacun des pixels la valeur moyenne correspondante des pixels des blocs voisins. De part son aspect récursif, ce type de correction ne correspond pas à la contrainte de faible coût.

En 2002, Gao et al [Gao et al., 2002] proposent de parcourir plusieurs fois les blocs pour détecter dans le cas des blocs étendus une zone dans laquelle le filtrage pourra être appliqué sans dégrader l'information. Cela permettra de filtrer sur plus de pixels, avec un noyau plus grand, réduisant ainsi le risque de frontières fantômes.

Sans utiliser d'autres blocs, Averbuch et al [Averbuch et al., 2005] proposent d'améliorer la correction dans le cas des zones fortement compressées en appliquant un filtre itératif pour corriger jusqu'à ce que la frontière ne se voit plus.

Ces méthodes sont capables de faire disparaître ou d'atténuer les frontières fantômes. Néanmoins, de part leur caractère itératif et récursif, ce type de correction est difficilement intégrable dans un algorithme de correction temps réel et c'est pourquoi le phénomène des frontières fantômes est encore largement présent pour des diffusions numériques à faible bande passante.

3.2 La confusion entre frontières de blocs et contours de l'image

Ce problème provient de la correction de l'effet de bloc sans l'utilisation des paramètres de quantification.

La cause La problématique se pose sur des images ayant des contours linéaires pouvant s'apparenter à des frontières de blocs. En effet, ces contours peuvent parfois venir se positionner sur la grille représentant le découpage de l'image en blocs. Dans le cas d'une compression de type MPEG-2 cette grille est de 8 pixels par 8 pixels. La figure 2.15 illustre une image type où les immeubles représentent des contours linéaires. Nous avons représenté en noir la position de la grille 8*8 délimitant les frontières de blocs. Nous voyons que les contours de l'image sont parfois positionnés sur cette grille.

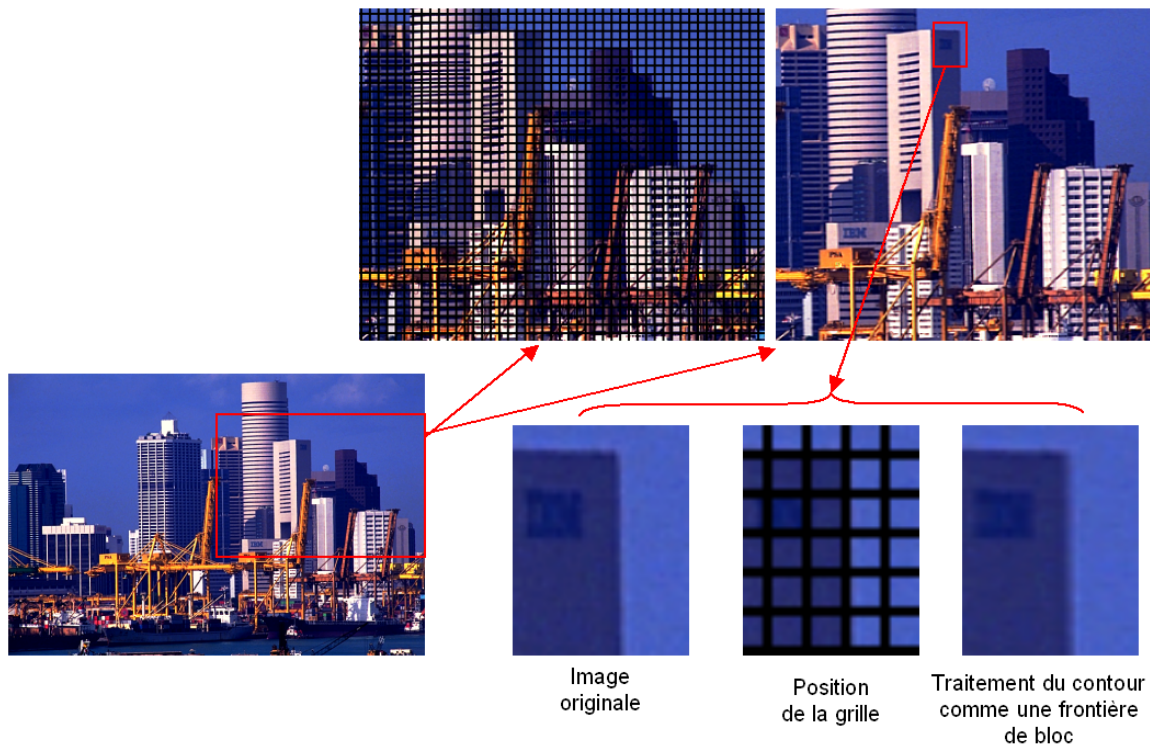


FIG. 2.15 – La confusion entre les frontières de blocs et les contours de l'image.

La conséquence Si on applique à ce genre d'image une correction de type classification et filtrage, la classification va confondre le contour de l'image avec une frontière de bloc et appliquer un filtrage passe-bas à cette position. Le résultat obtenu est représenté sur la figure 2.15 : le contour est filtré et peut être éliminé si le filtrage utilisé est trop agressif.

Les méthodes utilisées pour y remédier Ce type de problème peut être évité si l'on connaît le paramètre de quantification. En effet, un paramètre de quantification faible représentera une faible compression et donc une faible probabilité d'avoir une frontière de bloc aussi gênante. Inversement, un paramètre de quantification élevé symbolisera une forte compression et augmentera la probabilité que l'information rencontrée soit une frontière de bloc.

Remarque : Si l'information était effectivement un contour, la forte compression l'aurait endommagé et il n'apparaîtrait plus comme un contour net. La figure 2.16 montre de quelle manière est dégradé un contour à différents taux de compression : l'effet de ringing apparaît et le contour n'est plus net.



FIG. 2.16 – De gauche à droite : la zone de l'image est de plus en plus compressée et montre l'évolution de la dégradation du contour.

Pour certaines applications ayant accès aux paramètres du décodeur, ce coefficient de quantification est connu et aide à obtenir une correction plus robuste [Kaup, 1997] dans le sens où c'est une aide à la décision. En plus du paramètre de quantification, tous les paramètres issus du décodeur peuvent être utilisés.

Dans [Lee et Park, 2001], les auteurs proposent une correction post-traitement. En fonction du mode de codage incluant les coefficients DCT, le type de macrobloc et le paramètre de quantification, ils définissent un booléen pour les frontières de blocs horizontales, un booléen pour les frontières de blocs verticales et un booléen pour l'effet de ringing. Suivant l'état de ces booléens, ils appliquent différents types de filtres dans la boucle mais aussi en post-traitement. Ces différents types de filtrage sont résumés dans le tableau 2.1.

TAB. 2.1 – Les différents filtres proposés par Lee et al

	Dans la boucle	Post-Traitement
Fort Deblocking	Filtrage 1D sur 6 pixels	
Faible Deblocking	Filtrage 1D sur 2 pixels	Filtrage 1D sur 4 pixels
Ringing	Détection de contours et Filtrage 2D	

Dans le cas où ces paramètres ne sont pas disponibles, on peut supposer pouvoir les estimer. Chou et al [Chou et al., 1998] proposent une estimation du paramètre de quantification en utilisant des modèles probabilistes. Ces techniques sont malheureusement trop complexes pour notre

contrainte de faible coût.

Une autre possibilité est proposée par Kong et al [Kong et al., 2004] au travers d'une méthode globale de deblocking/deringing : leur méthode de distinction entre les frontières et les contours est basée sur des seuils trouvés expérimentalement (ils utilisent la valeur la plus fréquemment trouvée sur un panel d'images).

Le problème avec ce genre de solution expérimentale est que si le type d'image traité ne correspond pas aux types d'images expérimentales, le seuil utilisé n'est pas optimal et augmente le risque de dégradation.

Dans le cas des post-traitements n'ayant accès qu'aux images décodées et donc qu'aux valeurs des pixels, cette aide à la décision fournie par le paramètre de quantification n'est pas accessible. Or, les méthodes d'estimation sont soit trop coûteuses, soit pas assez robustes, c'est pourquoi le genre de dégradation, présentée sur la figure 2.15, est fréquemment rencontrée.

3.3 L'apparition du flou

Ce type de limitation est causé par la correction de l'effet de bloc et de l'effet de ringing et en l'absence des paramètres du décodeur.

Cause et conséquence liées à l'effet de bloc Le cas évoqué précédemment de confusion entre frontières et contours est un cas particulier de l'apparition du flou. En effet la détérioration du contour se traduit par son filtrage et c'est le filtrage de ces coefficients hautes fréquences qui provoque l'apparition du flou. On remarque d'ailleurs sur 2.15 que la dégradation causée par la confusion du contour et de la frontière est la génération du flou.

De plus, dans le cas où le contour n'est pas précisément sur la position de la grille 8x8 et que la taille du filtre est trop importante par rapport à l'information à traiter, le filtrage a lieu sur l'information contenue dans le bloc et causera aussi l'apparition du flou par l'élimination des informations du bloc.

Les méthodes utilisées pour y remédier Ces méthodes sont les mêmes que celles utilisées pour la confusion entre frontières de blocs et contours : elles s'appuient sur les paramètres du décodeur tels que le type de frontière (inter macrobloc ou intra macrobloc), le type d'images (Intra, Prédicative ou Bidirectionnelle) et le paramètre de quantification. Ces paramètres aident à la décision car ils renseignent sur le codage ou la prédiction du bloc et sur la force de compression. Ainsi, il est plus aisé de savoir si l'information a été endommagée, et si il faut la corriger évitant ainsi toute correction inutile pouvant causer une perte d'informations se traduisant par le phénomène de flou.

La cause liée à l'effet de ringing Dans le cas de la correction de l'effet de ringing, la première étape consiste à détecter les contours sur l'image entière ou sur un bloc. Puis, un filtre passe-bas vient supprimer le bruit le long des contours. Sur la figure 2.17, l'image A est l'image à traiter. L'image B est le résultat obtenu après l'application d'un filtre de Sobel pour détecter les frontières horizontales.

Nous avons déjà évoqué que l'inconvénient de ce genre de méthode est de donner un seuil à partir duquel l'information est considérée comme un contour. Sur l'image en bas à gauche, le seuil utilisé est égal à 100 : les contours du pied, des toits des immeubles sont représentés, mais

aussi des contours sur la pelouse qui ne sont pas très représentatifs et un filtrage passe-bas sur la pelouse causera l'apparition du flou. Inversement, si le seuil est fixé à 200, les plus faibles contours tels que les toits des immeubles ne sont pas tous pris en compte et le ringing proche de ces contours sera peu corrigé. La difficulté est donc de trouver le meilleur compromis qui est souvent dépendant du contenu de l'image.

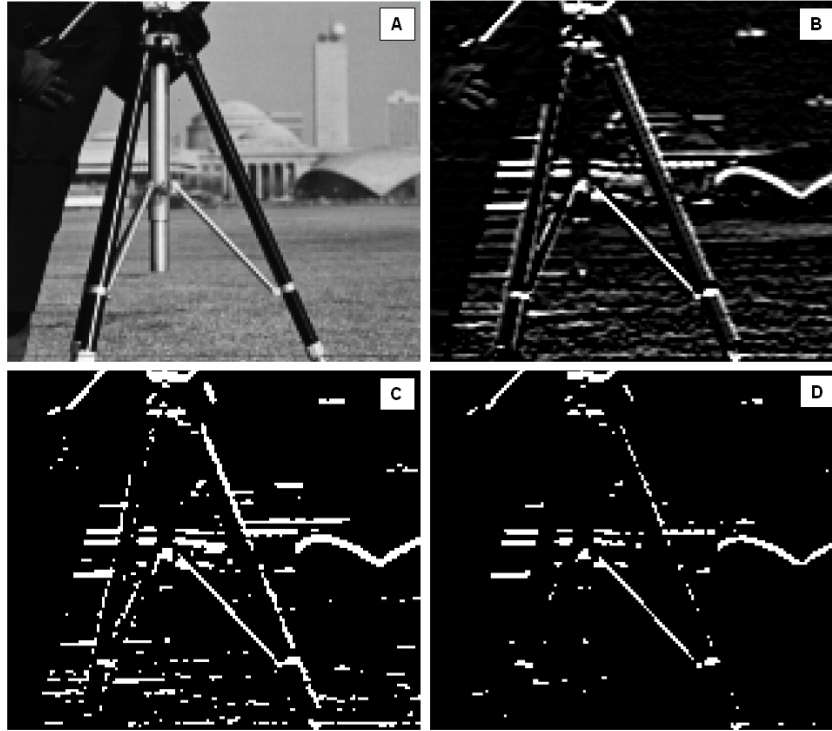


FIG. 2.17 – Problème du seuillage avec les filtres passe-haut. A : Image originale, B : Contours verticaux détectés par un filtre de Sobel. C : les intensités au dessus de 100 sont considérées comme des contours D : idem avec une valeur de seuil égale à 200.

Les méthodes utilisées pour y remédier Le réglage du seuil peut être automatique et s'adapter en fonction du contenu de l'image. Par exemple, les méthodes telles que le seuillage par maximisation de la variance interclasse [Otsu, 1979], par maximisation de l'entropie [Pun, 1981], ou encore par la méthode des nuées dynamiques [Diday, 1971] donnent de meilleurs résultats mais présentent un caractère itératif et nécessitent un traitement sur l'image entière les rendant inexploitable en temps réel. Pour ces raisons, le flou est donc encore fréquemment rencontré dans les corrections implémentées en temps réel.

Qu'il soit causé par la correction de l'effet de bloc ou la correction de l'effet de ringing, le flou peut être évité en tenant compte de plus de caractéristiques du système visuel humain. Par exemple, dans [Coudoux et al., 1998], les auteurs développent un modèle leur permettant d'agir plus précisément sur les zones de l'image en mouvement correspondant aux zones où le SVH est le plus fréquemment attiré. Puis, ils appliquent un filtre dont la réponse est directement liée à la visibilité des frontières de blocs.

Dans [Singh et al., 2007], les auteurs utilisent la fonction de sensibilité au contraste CSF et l'effet de masquage pour définir la visibilité d'une frontière. La fonction de sensibilité au contraste

modélise la réponse du SVH en fonction des différentes fréquences spatiales et permet de prédire dans quelles zones de luminance, le SVH est le plus sensible aux variations et donc aux frontières de blocs. Ces caractéristiques du système visuel humain sont souvent prises en compte dans le domaine fréquentiel [Liu et Bovik, 2002] ou spatial [Shi et al., 2006]. En effet, ces caractéristiques permettent d'éviter une étape de correction dans les zones où la frontière existe mais ne se voit pas.

Remarque : en revanche, ce genre de paramètre de décision n'est pas utilisé dans les algorithmes temps réel dans la mesure où il est moins coûteux d'appliquer une correction qui ne se verra pas plutôt que d'effectuer une étape de condition puis de correction. Néanmoins, il faut reconnaître que ce type de méthode permet de ne pas corriger à certains endroits évitant ainsi de faire une mauvaise correction et de dégrader l'image.

Même si les techniques de classification par bloc et filtrage sont les plus adaptées à notre contexte d'étude, la correction de l'effet de bloc et de l'effet de ringing peut provoquer l'apparition d'autres artefacts tels que les frontières fantômes, la dégradation des contours et l'apparition du flou. Ces artefacts sont répertoriés sur la figure 2.18.

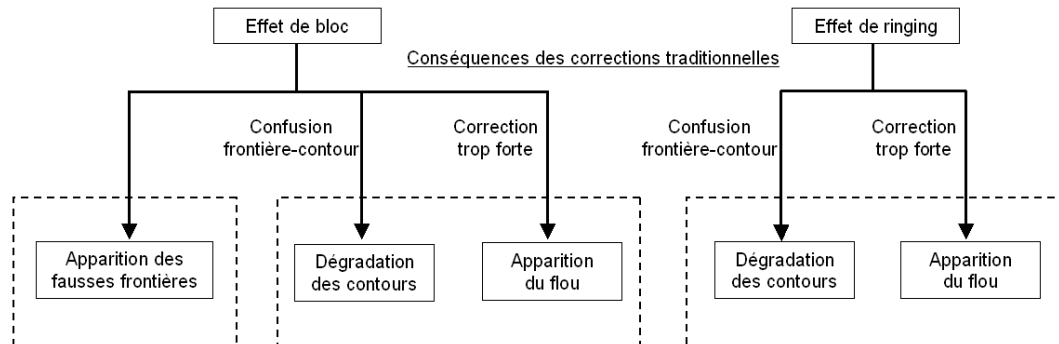


FIG. 2.18 – Les artefacts de correction : conséquences du traitement des artefacts de correction.

Dans ce chapitre, nous avons établi un état de l'art des corrections proposées pour supprimer les artefacts de compression. Nous avons ciblé l'effet de bloc et l'effet de ringing car ils sont considérés comme les plus gênants et leurs corrections peuvent atténuer la visibilité des autres artefacts.

Au travers de cet état de l'art, nous avons différencié les méthodes de pré-traitement, les méthodes d'amélioration du principe de l'encodeur, les méthodes de correction dans la boucle et les méthodes de correction sur l'image décodée. Pour être applicable dans notre contexte, nous avons sélectionné les méthodes de corrections de l'image en post-traitement. Cherchant à obtenir la meilleure qualité d'image, les méthodes d'amélioration étaient préférables aux méthodes de restauration qui de toutes façons étaient difficilement implémentables en temps réel.

Puis, pour répondre à nos contraintes d'implémentation notre choix s'est tourné vers l'étude précise des méthodes de corrections par classification du bloc suivi d'un filtrage. L'analyse de la position de la frontière et de l'information au sein des blocs se fait dans le domaine spatial et non dans le domaine fréquentiel évitant ainsi l'estimation des coefficients DCT. Ces méthodes sont très appropriées aux corrections de l'effet de bloc et du ringing mais peuvent vite générer d'autres artefacts si elles sont mal contrôlées et mal utilisées.

Mettant en évidence les artefacts tels que les fausses frontières ou le flou causés par ce type de correction, nous soulignons l'importance d'en tenir compte pour estimer la qualité d'une correction. La simple suppression d'un artefact n'implique pas forcément une bonne correction. L'évaluation de l'efficacité d'une correction va donc se traduire par sa capacité à réduire les artefacts de compression sans en faire apparaître d'autres.

Bilan de la première partie

Cette première partie était consacrée à l'analyse du sujet et à la mise en évidence de la problématique. Les différentes étapes sont représentées sur la figure 2.19.

Après avoir replacé la compression MPEG-2 dans l'historique des normes, nous avons étudié plus précisément les techniques utilisées pour la compression. Malgré les efforts pris en compte par les développeurs des normes pour obtenir une image de bonne qualité visuelle, la limitation de la bande passante est parfois telle que les dégradations sont trop importantes et deviennent perceptibles et gênantes pour le SVH.

Nous avons décrit ces artefacts en expliquant les causes et les conséquences de chacun sur notre perception d'une image. Parmi ces artefacts, l'effet de bloc et l'effet de ringing nous ont paru prioritaires. Pour ces artefacts, la littérature propose un panel important de méthodes capables d'atténuer ou de supprimer ces artefacts. Dans l'optique de les implémenter dans notre contexte d'étude avec des contraintes définies en introduction, les méthodes de classification et filtrage dans le domaine spatial nous ont paru les plus adaptées. Ces méthodes déjà très étudiées, sont efficaces mais peuvent être responsables de l'apparition d'autres dégradations que nous avons appelées les artefacts de correction.

Cette étude nous a permis de mettre en évidence la problématique du travail demandé sur la réduction des artefacts de compression MPEG-2 : l'efficacité d'une correction est non seulement liée à sa capacité à supprimer les artefacts de compression mais aussi à sa capacité à ne pas faire apparaître des artefacts de correction.

Pour contrôler ce phénomène, nous proposons d'utiliser des métriques évaluant la visibilité de ces artefacts pré et post-correction. Ces métriques nous permettront d'évaluer la qualité des différentes corrections et nous guideront sur nos choix d'éventuelles améliorations.

L'analyse et l'utilisation de ces métriques font l'objet de la seconde partie de ce travail.

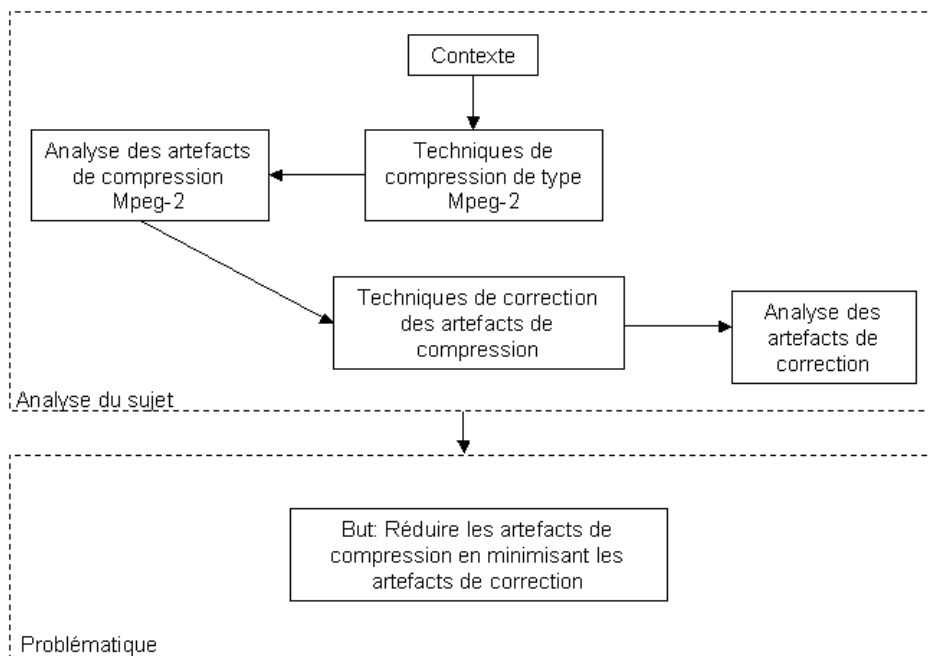


FIG. 2.19 – Bilan de la première partie.

Deuxième partie

La mesure des artefacts

Chapitre 3

L'estimation de la qualité

La perception des artefacts dépend de facteurs externes. Ces facteurs peuvent être liés aux caractéristiques du SVH (les artefacts se distinguent moins bien dans les zones texturées que dans les zones homogènes) mais aussi aux caractéristiques d'observations (plus l'observateur se positionnera loin de l'écran, moins il portera d'attention aux effets de flou). Partant de ces remarques, comment évaluer les artefacts perçus par un utilisateur dans des conditions précises de visualisation ?

Cette question est légitime dans la mesure où la qualité d'images est de plus en plus recherchée notamment lorsqu'on s'attache au marché de la télévision numérique qui, pour tout utilisateur doit prétendre à une qualité supérieure sinon au moins égale à la qualité des chaînes analogiques.

Le premier objectif de ce chapitre est de définir les notions de qualité et de fidélité. A partir de ces deux notions, nous expliquerons les protocoles utilisés pour estimer une note représentative de l'avis de la population. Mettant en évidence les contraintes de temps, de gestions et d'analyses liées à ces protocoles, nous montrerons comment des mesures automatiques de fidélité et de qualité peuvent les remplacer dans des conditions bien déterminées.

1 Bien définir le terme de qualité

La perception désigne la sensation éprouvée par l'observateur à qui on présente un stimulus. C'est cette perception qui va lui permettre de porter un jugement positif ou négatif. Ce jugement peut porter sur l'ensemble de l'image ou de la vidéo, ou plutôt s'axer sur une caractéristique précise. De plus, ce jugement peut être issu d'une comparaison entre deux données. Ces quelques remarques montrent la complexité du terme qualité et l'intérêt de définir correctement certaines notions.

1.1 La notion de fidélité

Pour une image, la notion de fidélité permet de savoir si la reproduction d'une image est fidèle ou non à l'original. La notion de fidélité repose sur les différences pixels à pixels établies entre ces deux images. Une image peut être fidèle mais de mauvaise qualité et une image de qualité correcte peut parfois avoir une fidélité assez mauvaise.

1.2 La notion de qualité

La mesure de qualité, en revanche, cherche à quantifier l'appréciation de l'observateur selon sa conscience, ce qui la rend très difficile à évaluer.

Remarque : Cette différence de notion entre fidélité et qualité peut respectivement être mise en parallèle avec la différence entre les notions de méthodes de reconstruction et d'amélioration d'images.

La qualité peut être subjective ou objective.

La qualité subjective Elle s'obtient au travers de tests subjectifs dont le processus est le suivant : un échantillon représentatif de la population juge de la qualité d'une série d'images ou de vidéos. Ces jugements sont récupérés et analysés pour obtenir une note de qualité moyenne. La qualité subjective représente la vérité du terrain.

La qualité objective Elle consiste à considérer la mesure de la qualité comme une évaluation quantitative de la perception des défauts et de la gêne qu'ils procurent sur le traitement des informations perceptuelles [Barba, 1981]. Pour se différencier des mesures de fidélité, la mesure de qualité peut tenir compte des caractéristiques du système visuel humain. Pour être validée, une mesure objective de la qualité doit se rapprocher le plus possible des mesures de qualité subjectives obtenues pour un ensemble d'observateurs. Nous proposons de développer ces deux aspects de la qualité.

En utilisant ces définitions, nous allons décrire les méthodes subjectives, les méthodes objectives de fidélité et les méthodes objectives de qualité en insistant sur les points forts et les points faibles de chacune.

2 Les mesures subjectives

La qualité subjective s'obtient au travers de tests subjectifs. Nous commencerons par définir l'environnement utilisé pour la réalisation de ces tests. Puis, nous verrons comment les données récoltées par ces tests subjectifs sont transformées en une note de qualité finale : le "Mean Opinion Score" ou MOS.

2.1 Un protocole bien défini

Dans [Charrier et al., 2005], les auteurs proposent un bon résumé des conditions essentielles de normalisation. Elles sont obtenues en concordance avec les standards de l'ISO [ISO, 2000] et de l'ITU [BT.500-11, 2002]. Nous proposons de lister quelques-unes de ces caractéristiques.

2.1.1 Les conditions de visualisation

Ces recommandations sont principalement matérielles.

- Pour contrôler les conditions de visualisation, l'éclairage de la salle (tubes fluorescents D65 d'intensité variable contrôlée) doit être la seule source de lumière.
- La position et la distance de l'observateur par rapport à l'écran doivent être optimisées pour que le champ visuel de l'observateur ne soit pas gêné.
- L'écran doit être calibré en terme de luminosité maximale, de correction gamma et de température de couleurs.
- Les séances ne doivent pas dépasser une demi-heure pour éviter la fatigue et donc éviter de parasiter les résultats. Ces séances doivent débuter par une série de présentations non prises en compte dans les résultats finaux pour stabiliser l'opinion de l'observateur. Les expériences se font toujours en deux parties pour vérifier la cohérence des réponses.

2.1.2 Les observateurs

Tout comme les conditions d'observations, l'observateur doit lui aussi être 'normalisé'. En effet, le panel d'observateurs doit représenter un échantillon moyen de la population concernée par la nature de l'étude à mener. De plus, ces personnes sont généralement sélectionnées suivant leur acuité visuelle et leur vision normale des couleurs respectivement vérifiées par le test de Snellen et le test d'Ichihara représentés sur la figure 3.1.

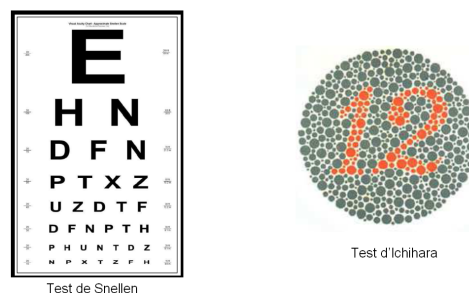


FIG. 3.1 – Les tests de Snellen et d'Ichihara.

2.1.3 Le choix des images et des séquences

La base de données d'images ou de séquences doit être en relation avec le thème du test. Le test se déroule généralement en deux parties pendant lesquelles la base de données est visualisée dans un ordre aléatoire. Cela permet d'étudier la cohérence des résultats pour un même observateur.

2.1.4 La précision des directives du test

Outre la rigueur matérielle, ces tests nécessitent une grande précision dans leur énoncé. En effet, ce que l'observateur doit juger doit être clairement défini pour éviter toute confusion ou pour éviter de fausser les résultats.

De plus, le protocole doit être adapté en fonction de l'application. Par exemple, pour juger de la qualité d'images médicales, seules des personnes qualifiées seront aptes à comprendre ces images et donc à les noter. Dans ce cas, l'échantillon moyen de la population utilisé pour les tests subjectifs sera restreint à des personnes du monde médical. En revanche, pour évaluer la qualité des images de la télévision, l'échantillon moyen sera beaucoup plus large en terme de sexe, d'âge et d'origine sociale.

Malgré toutes ces précautions prises pour effectuer des tests dans les conditions les plus neutres possibles, le vécu de chacun joue malgré tout un rôle dans la note finale. Dans ce cas, plus le nombre d'échantillons de personnes sera élevé, moins cet aspect apparaîtra dans les résultats finaux. L'ITU recommande au moins 15 observateurs.

2.2 Les tests psychophysiques sur les images fixes

Les normes internationales proposent deux familles de tests psychophysiques pour les images fixes : les tests comparatifs qui permettent de noter une image en fonction d'une autre et les tests de mesure absolu qui permettent de noter une image sur sa qualité sans aucun moyen de comparaison.

2.2.1 Les tests comparatifs

Pour ces tests, les utilisateurs doivent comparer plusieurs images les unes par rapport aux autres. Par exemple, le test d'ordonnement (fig. 3.2) permet à l'observateur de classer une série de 9 images de la meilleure qualité à la plus mauvaise, lui attribuant ainsi une note de 1 à 9. Il existe également le test de choix forcé qui consiste à comparer les images deux à deux.

Notons que ces tests ne permettent pas de dire qu'une image est de bonne qualité mais qu'elle est meilleure qu'une autre. Ce genre de test peut être utilisé pour comparer l'efficacité d'un algorithme de correction par rapport à d'autres. Néanmoins, l'algorithme classé en première position ne sera pas pour autant de très bonne qualité. Le test ne dit pas non plus si les différences sont significatives ou pas.



FIG. 3.2 – Le test d’ordonnement.

2.2.2 Les tests de mesures absolues

Le but de ces tests est de quantifier numériquement la qualité de l’image affichée à l’écran suivant un type d’échelle de notation représentée sur le tableau suivant 3.1.

TAB. 3.1 – Echelle de notation.

Qualité	Note	Dégradation
Excellente	5	Imperceptible
Bonne	4	Perceptible mais non gênant
Assez bonne	3	Légèrement gênant
Médiocre	2	Gênant
Mauvaise	1	Très gênant

Contrairement aux tests comparatifs, c’est bien la qualité qui est notée suivant le degré de perceptibilité de la dégradation.

Remarque : L’ordre dans lequel la qualité des images est estimée, même si il est aléatoire, joue un rôle dans la perception de la qualité. Par exemple, une même image de mauvaise qualité risque d’être notée plus sévèrement si elle est visualisée après une image de mauvaise qualité. Même si ce n’est pas un test de comparaison en soi, le SVH a malgré tout tendance à se calibrer par rapport à ce qu’il vient de voir.

2.3 Les tests psychophysiques sur les séquences animées

Pour les séquences, 3 types de tests sont préconisés : les méthodes à simple stimulus, les méthodes à double stimulus et les méthodes comparatives.

2.3.1 Les méthodes à simple stimulus

Lorsque l'on ne dispose pas des séquences originales, les méthodes à simple stimulus ou "Single Stimulus Continuous Quality Scale" (SSCQS) sont utilisées. Les séquences sont présentées une par une comme cela est schématisé sur la figure 3.3 et évaluées indépendamment sur une paire d'échelles verticales représentées sur la figure 3.4.

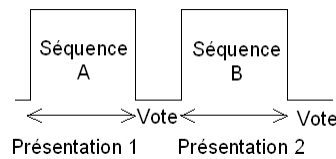


FIG. 3.3 – Le protocole des méthodes à simple stimulus.

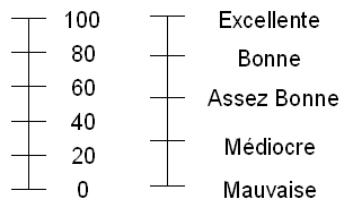


FIG. 3.4 – A gauche : échelle de notation continue. A droite : échelle de notation par catégories de qualité.

2.3.2 Les méthodes à double stimuli

La méthode à double stimuli, appelée aussi DSCQS (Double Stimuli Continuous Quality Scale), est dédiée aux tests nécessitant la séquence originale. Une paire de séquences vidéo (séquence de référence et séquence dégradée) est présentée à l'observateur qui doit porter son jugement suivant les échelles verticales continues utilisées pour la méthode à simple stimuli (fig. 3.4). La figure 3.5 schématise un exemple des phases se déroulant dans ce test : la présentation contenant la vidéo de référence et la vidéo à évaluer, le vote, et ainsi de suite sur le nombre de séquences présentées. Les séquences peuvent être visualisées en même temps ou l'une après l'autre, la position de la vidéo de référence étant aléatoire.

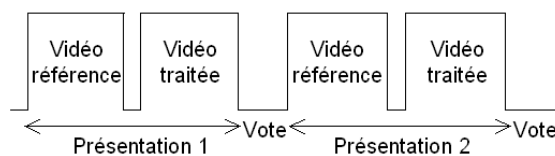


FIG. 3.5 – Le protocole des méthodes à double stimuli.

2.3.3 Les méthodes comparatives

Pour ces méthodes, une paire de séquences vidéo est également présentée aux observateurs. L'originalité provient de son échelle de notation comparative fournie par l'ITU qui est représentée sur la figure 3.6. Ce type d'échelle ne permet pas d'affirmer que la meilleure des vidéos est de bonne qualité sauf si il y a un moyen d'être sûr que la vidéo de référence est de bonne qualité.

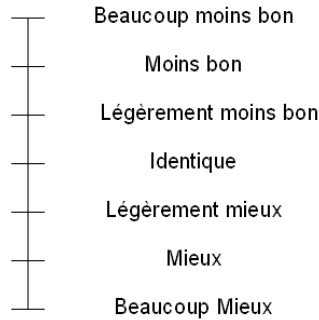


FIG. 3.6 – L'échelle comparative.

Remarque : Ces méthodes (simple stimulus, double stimuli, comparatives) s'appliquent aussi aux images statiques. Néanmoins, il existe une méthode plus particulièrement dédiée aux séquences vidéos : la méthode SAMVIQ pour "Subjective Assessment Methodology for Video Quality". Nous n'utiliserons pas cette méthode destinée aux contenus multimédia et nous laisserons le lecteur se référer au document de l'ITU [ITU, 2004].

Bilan L'intérêt de ces différentes échelles et de ces tests est de pouvoir juger la dégradation des images ou des vidéos de manières différentes. Ces tests donnent un ensemble de note qui vont être traitées de manière à obtenir une note finale.

2.4 La note objective de qualité

2.4.1 Le calcul de l'opinion moyenne : le "Mean Opinion Score"

La première étape de l'analyse des résultats consiste à calculer le Mean Opinion Score, c'est-à-dire la note d'opinion moyenne dont la formule est la suivante :

$$MOS_{jk} = \frac{1}{N_{obs}} \sum_{i=1}^{N_{obs}} Nsub_{jk}(i, l) \quad (3.1)$$

où N_{obs} est le nombre d'observateurs ayant noté l'image originale j subissant la k^{ime} dégradation et $Nsub_{jk}(i, l)$ la note fournie par l'observateur i ayant noté l'image originale j subissant la k^{ime} dégradation

2.4.2 L'intervalle de confiance

Pour évaluer au mieux la fiabilité des résultats, il est préférable d'associer à chaque MOS un intervalle de confiance. En général, on utilise l'intervalle de confiance à 95%.

$$[MOS_{jk} - e_{jk}, MOS_{jk} + e_{jk}]$$

où $e_{jk} = 1.96 * \frac{\sigma_{jk}}{\sqrt{n}}$ et

$$\sigma_{jk} = \sqrt{\sum_{i=1}^{N_{Obs}} \frac{(N_{sub_{jk}}(i) - MOS_{jk})^2}{N_{Obs} - 1}} \quad (3.2)$$

2.4.3 L'analyse statistique des résultats

Dans certains tests, la même image peut être jugée plusieurs fois par l'observateur. Ce test permet d'analyser la cohérence des résultats d'un même observateur. Ainsi, l'ITU recommande parfois le rejet de certains résultats. Ce mode de sélection dépend du type de test subjectif effectué. Pour une échelle allant de 1 à 5, si pour une même image jugée dans les mêmes conditions par le même observateur, les notes diffèrent de 2 points ou plus, ces notes sont rejetées.

Puis, la cohérence des notes d'un observateur avec les notes moyennes est étudiée au travers du test du β_2 (expliqué dans [BT.500-11, 2002]).

Enfin, un observateur peut également, sur l'ensemble de ces jugements, surévaluer ou sous-évaluer la qualité par rapport aux autres observateurs. Dans ce cas, l'utilisation de la transformation inverse des Z-scores expliquée dans [Charrier et al., 2005] permettra de redonner les notes subjectives de qualité comprises dans l'échelle de notation.

Ces tests subjectifs donnent une note de qualité représentative de ce que doit percevoir globalement tout observateur concerné par l'application. Ces tests, utilisés pour quantifier la qualité d'une image ou d'une vidéo, sont malheureusement soumis à plusieurs contraintes que nous allons lister.

2.5 Les contraintes des tests subjectifs

Ces contraintes peuvent être classées en 3 catégories qui sont la gestion, le temps et l'analyse.

2.5.1 Les contraintes de gestion

Ces tests nécessitent :

- l'accès à une salle dont les caractéristiques se rapprochent le plus possible de celles données par l'ITU.
- un nombre de personnes suffisant et représentatif de la population, ce qui se trouve pour un test mais qui devient difficile pour plusieurs séries de tests.
- une ou plusieurs personnes qui coordonnent ces tests.
- une base de données d'images adéquates à l'application ciblée.

2.5.2 Les contraintes de temps

Dans le meilleur des cas, une série de tests subjectifs prendrait une durée minimum égale à la durée du test par le nombre d'observateurs, soit en moyenne une trentaine de minutes pour une vingtaine d'observateurs, c'est-à-dire une dizaine d'heures. Bien sur, ce nombre peut être réduit si on fait passer plusieurs tests en même temps mais, dans ce cas, les contraintes de gestion

augmentent : il faut plusieurs salles, plusieurs écrans calibrés de la même façon et plusieurs personnes pour coordonner ces tests.

2.5.3 Les contraintes d'analyses

La vérification de l'acuité visuelle de chaque observateur, l'analyse statistique des données et le calcul du MOS montrent la difficulté et la rigueur d'analyse nécessaires pour utiliser les résultats obtenus. De plus, l'interprétation des résultats en fonction de l'application visée est délicate et dépend de la base de donnée sélectionnée.

L'avantage de ces tests subjectifs est la validité des résultats obtenus puisque c'est le SVH qui juge. Néanmoins, ces résultats peuvent facilement être faussés de part le non-respect des contraintes de temps, de gestion et d'analyse. De plus, ces tests sont difficilement réalisables pour analyser les images diffusées. Nous allons donc étudier les mesures objectives qui donneront automatiquement une note. Ces mesures objectives peuvent être de fidélité ou de qualité.

3 Les mesures objectives de fidélité

Les premières mesures objectives développées ont été des mesures de fidélité. Ces mesures peuvent être des simples calculs de distance ou des mesures plus élaborées utilisant les critères du traitement du signal.

3.1 Les mesures de distances

Nous présentons ici des critères de mesures de distance fondés sur la norme L_p définie par :

$$L_p \{F, \tilde{F}\} = \left\{ \frac{1}{M * N} \sum_{j=1}^M \sum_{k=1}^N |F(j, k) - \tilde{F}(j, k)|^p \right\}^{1/p} \quad \text{pour } p \geq 1 \quad (3.3)$$

où F est l'image originale, \tilde{F} l'image dégradée, M et N la hauteur et la largeur de l'image. Pour une valeur importante de p , les erreurs les plus importantes seront prises en compte.

A partir des différentes valeurs de p , nous obtenons :

$$L_1 \{F, \tilde{F}\} = AD = \left\{ \frac{1}{M * N} \sum_{j=1}^M \sum_{k=1}^N |F(j, k) - \tilde{F}(j, k)| \right\} \quad \text{pour } p = 1 \quad (3.4)$$

c'est-à-dire la différence moyenne (Average Difference) entre les images F et \tilde{F} .

$$L_2 \{F, \tilde{F}\} = RMSE = \left\{ \frac{1}{M * N} \sum_{j=1}^M \sum_{k=1}^N |F(j, k) - \tilde{F}(j, k)|^2 \right\}^{1/2} \quad \text{pour } p = 2 \quad (3.5)$$

qui est la racine de l'erreur quadratique moyenne ou RMSE (Root Mean Square Error). En pratique, on utilise L_2^2 plutôt que L^2 , c'est-à-dire l'erreur quadratique moyenne MSE (Mean Squared Error).

$$L_2^2 \{F, \tilde{F}\} = MSE = \frac{1}{M * N} \sum_{j=1}^M \sum_{k=1}^N |F(j, k) - \tilde{F}(j, k)|^2 \quad (3.6)$$

Ces deux mesures ont les mêmes propriétés pour leurs minima et maxima, mais la MSE est plus sensible aux grandes différences que la RMSE.

La MSE peut être présentée sous une forme normalisée :

$$PMSE = \left\{ \frac{1}{M * N} \sum_{j=1}^M \sum_{k=1}^N |F(j, k) - \tilde{F}(j, k)|^2 \right\} / \{Max F(j, k)\}^2 \quad (3.7)$$

qui est le "Peak" MSE, ou encore

$$NMSE = \left\{ \sum_{j=1}^M \sum_{k=1}^N |F(j, k) - \tilde{F}(j, k)|^2 \right\} / \left\{ \sum_{j=1}^M \sum_{k=1}^N (F(j, k))^2 \right\} \quad (3.8)$$

Ces normalisations rendent ces mesures de distance moins dépendantes de l'image de référence.

3.2 Les critères issus du traitement du signal

Le critère le plus couramment utilisé dans la littérature pour spécifier la qualité d'une image est le PSNR (Peak Signal Noise Ratio). Si on considère que l'image originale est un signal d'entrée $f(j,k)$ et que l'image dégradée correspond au signal de sortie du système $g(j,k)$, alors l'image de différence peut être considérée comme étant un bruit $e(j,k)$, c'est-à-dire :

$$g(j, k) = f(j, k) + e(j, k) \quad (3.9)$$

Il devient alors possible d'exprimer le rapport signal à bruit (Signal Noise Ratio) d'un tel système :

$$SNR = 10 \log_{10} \frac{\sum_{j=1}^M \sum_{k=1}^N [g(j, k)]^2}{\sum_{j=1}^M \sum_{k=1}^N [e(j, k)]^2} (dB) \quad (3.10)$$

Ce rapport peut être exprimé d'une autre manière en utilisant une mesure de type MSE par :

$$SNR = 10 \log_{10} \frac{\frac{1}{MN} \sum_{j=1}^M \sum_{k=1}^N [g(j, k)]^2}{MSE} (dB) \quad (3.11)$$

Ainsi, on peut définir le PSNR :

$$PSNR = 10 \log_{10} \frac{(\text{valeur max du signal})^2}{MSE} = 10 \log_{10} \frac{255^2}{MSE} (dB) \quad (3.12)$$

Ces mesures de distances, très simples à mettre en oeuvre présentent quelques limitations que nous allons présenter.

3.3 La problématique de ces mesures de fidélité

Ces mesures sont dites de fidélité car elles estiment si la reconstruction est fidèle ou non. Depuis quelques années, des études sont faites pour insister sur le fait que ces métriques ne peuvent pas être utilisées en terme de qualité proprement dite. En effet, dans [Girod, 1993], [Eskicioglu, 2000], et [Wang et Bovik, 2006], il est bien mis en évidence que pour une même valeur de MSE ou de PSNR, on peut avoir des qualités d'images totalement différentes.

La figure 3.7 représente une série de 3 images : l'image originale, l'image originale dégradée par du bruit "sel et poivre" et l'image originale floutée. Le MSE et le PSNR obtenus sont très proches pour les deux images dégradées alors que le niveau de dégradation est différent. Ce phénomène s'explique par le fait que le SVH est plus sensible à un changement de structures dans l'image plutôt qu'à un changement pixel à pixel. Malgré leurs défauts, le MSE et le PSNR restent très employés car ils sont simples à calculer et facilement modélisables mathématiquement. Cependant ils nécessitent l'accès à l'image originale ce qui limite le type d'application.



FIG. 3.7 – Défaut du PSNR et du MSE dans l'estimation de la qualité. Par rapport à l'image originale, les deux images bruitées ont une qualité perçue différente mais un PSNR et un MSE très proche.

Remarque : Ces mesures ne tiennent pas compte de l'aspect couleur. Elles peuvent être appliquées composante par composante, quelque soit l'espace, mais devront être interprétées comme une mesure de fidélité associée à cette composante. Nous reviendrons à une mesure spécifique de qualité liée à la couleur dans le chapitre 5 section 4.

Nous venons de décrire les mesures objectives de fidélité et pourquoi elles ne pourront pas nous aider à estimer la visibilité des artefacts de compression et des artefacts de correction. Nous allons donc étudier les mesures de qualité objectives.

4 Les mesures de qualité objectives

Nous commencerons par définir les différentes catégories de métriques objectives de qualité. Puis, nous verrons les étapes nécessaires à sa validation. Enfin, nous verrons qu'elles peuvent estimer la qualité générale ou un artefact précis. Suite à ces descriptions, nous reviendrons à notre contexte d'étude pour définir le type de métriques de qualité dont nous avons besoin.

4.1 Les différentes catégories de métriques de qualité objectives

Le but d'une métrique de qualité objective est de donner automatiquement sur une échelle définie, la note de qualité la plus proche possible de ce qu'aurait jugé un observateur. C'est pourquoi, certaines métriques utilisent les caractéristiques du SVH ou encore des informations sur l'image ou sur le décodeur. Partant de ces informations, les métriques peuvent être développées suivant deux méthodes différentes. Après avoir défini toutes ces notions, nous illustrerons nos explications par deux exemples.

4.1.1 Les caractéristiques du système visuel humain

La perception d'une image peut se décomposer en trois phases : sa visualisation, sa transmission et son interprétation. Le système visuel humain a fait l'objet de nombreux travaux et de nombreuses expériences. Les caractéristiques issues de ces expériences se classent en deux catégories :

- Les caractéristiques bas niveau prenant en compte les deux premières étapes : visualisation et transmission
- Les caractéristiques haut niveau tenant compte de l'interprétation de ce que le SVH capte et de ce que le circuit neuronal lui transmet. Nous ne détaillerons pas ces caractéristiques qui sont liées à l'aspect psychophysique de l'observateur et qui seront donc propres à chacun. Elles sont trop difficiles à inclure dans un modèle de système visuel.

Parmi les caractéristiques bas niveau, nous commencerons par la capacité du SVH à détecter des différences de luminance, ce qui nous amène à définir la notion de contraste.

Le contraste représente le rapport sans dimension entre la luminance locale et la luminance moyenne. Suivant les expériences réalisées et les stimuli utilisés, plusieurs définitions du contraste ont été proposées.

- Le contraste de Michelson :

La luminance perçue par l'observateur humain dépend de la luminance de son voisinage. Notre perception est alors plus sensible à un contraste de luminance qu'à la valeur absolue de la luminance. Le contraste d'un motif périodique tel qu'une mire sinusoïdale peut être donnée par la formule de Michelson [Michelson, 1927] :

$$C_{Michelson} = \frac{L_{max} - L_{min}}{L_{max} + L_{min}} \quad (3.13)$$

où L_{max} et L_{min} désignent respectivement les valeurs maximales et minimales de luminance.

- Le contraste de Weber :

L'expérience de Weber [Peli, 1990] donne la définition suivante du contraste : soit un champ visuel de luminance uniforme L divisé en deux selon une ligne, l'observateur peut augmenter l'une des deux parties par une quantité variable ΔL . L'observateur ne remarque visuellement cette variation que lorsque ΔL atteint une valeur précise non nulle. Cette valeur est le seuil de détection en luminance ou JND ('Just Noticeable Difference'). En répétant cette expérience pour plusieurs valeurs de L , on remarque que le rapport $\Delta L/L$ est constant :

$$\frac{\Delta L}{L} = C_{Weber} \quad (3.14)$$

Ces définitions sont données dans le cas de stimuli simples. Pour des stimuli plus complexes, d'autres définitions sont données, notamment la définition par Peli [Peli, 1990] d'un contraste local.

La fonction de sensibilité au contraste CSF Elle va permettre de modéliser la réponse du SVH en fonction des différentes fréquences spatiales. La figure 3.8 représente la variation de la sensibilité au contraste en fonction des fréquences. On remarque le caractère passe-bas de cette fonction.

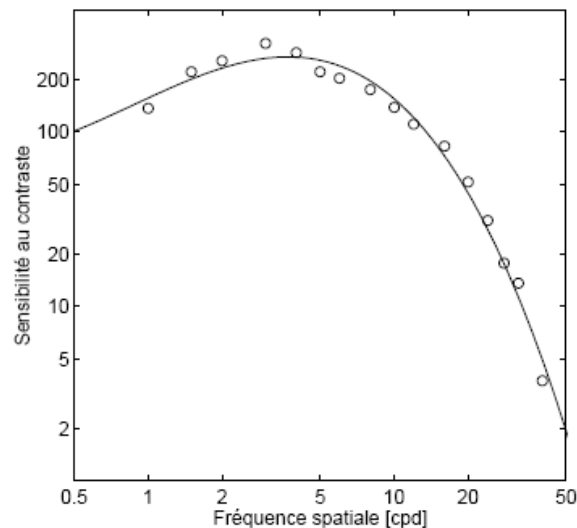


FIG. 3.8 – La sensibilité au contraste en fonction de la fréquence spatiale selon [Van Nes et Bouman, 1967], cpd : cycles par degré.

Les phénomènes de masquage Des expériences psychophysiques [Nachmias et al., 1971] mettent en évidence la structure multi-résolution du système visuel humain. Ceci a permis de donner les premières définitions des effets de masquage. On parle d'effet de masquage lorsqu'un signal peut modifier la visibilité d'un autre signal. La sensibilité au contraste représente déjà l'aspect de masquage de contraste et est l'une des causes du masquage spatial.

Le masquage spatial Il est directement lié à la CSF. En effet, le SVH étant moins sensible dans les hautes fréquences, il sera moins sensible au bruit dans les zones avec du détail.

L'adaptation rétinienne Ce phénomène peut être décrit par l'expérience suivante : lorsque nous passons d'un endroit fortement éclairé à un endroit sombre, notre perception initiale est fortement limitée mais avec le temps, elle s'accommode au faible éclairage. Inversement, si on passe d'un endroit sombre à un endroit éclairé. Ce phénomène traduit également un effet de masquage mais qui est temporel et qui va masquer les informations, le temps que le SVH s'accommode.

Le masquage temporel Au même titre que pour les fréquences spatiales, le SVH va être sensible à certaines propriétés des fréquences temporelles. Tout d'abord, la persistance rétinienne, valable pour l'accommodation du SVH aux images fixes, est également valable pour les séquences vidéos. En effet, si un changement brusque de luminosité intervient dans la séquence, le SVH, pendant le temps de l'accommodation, ne va pas percevoir toutes les informations des images qui défilent.

Ce phénomène se retrouve également après un changement de scène. Tam *et al* [Tam et al., 1995] ont examiné la visibilité des artefacts de codage après un changement de scène et il est apparu un phénomène de masquage significatif au niveau de la première image après le changement de scène. Si une image initialement dégradée intervient ponctuellement dans une séquence vidéo (pour une séquence de fréquence 25Hz, l'image apparaît 1/40 s), elle ne sera donc pas perçue. Ceci n'exclut pas le fait que les artefacts propres aux images fixes se retrouvent dans les séquences.

Enfin, le masquage temporel peut être dû à un objet en mouvement. Si l'attention est focalisée sur un objet en mouvement, les autres objets seront partiellement masqués. Ce phénomène est difficile à modéliser car comme il dépend de l'attention que va porter un observateur à une scène, il sera propre à son vécu et à ses préférences. Néanmoins, certaines études sont axées sur la recherche des points caractéristiques qui attirent notre attention. Les travaux de [Guironnet, 2006] proposent une méthode de résumé à partir de l'attention visuelle sur des séquences animées.

Bilan Ces caractéristiques bas niveaux sont plus ou moins prises en compte dans les métriques objectives. Cela dépend non seulement de l'application, image fixe ou séquence d'images, mais aussi de la possibilité d'intégrer tous ces phénomènes. De plus, ces caractéristiques sont souvent traduites pour un environnement et des conditions bien définies qu'il est difficile de généraliser.

4.1.2 Les différents accès à l'image non dégradée

Pour compléter ces phénomènes bas niveaux, les métriques peuvent s'appuyer sur différents niveaux de connaissances de la source non dégradée. Les métriques objectives peuvent être de 3 catégories qui vont dépendre du niveau d'informations accessibles sur la source et sur les traitements qu'elle a subis.

Les métriques avec référence Ces métriques utilisent l'image ou la vidéo de référence pour calculer la note de qualité. C'est une aide non négligeable qui permet d'avoir accès au contenu initial de l'image et de comparer ce qui a été ajouté ou supprimé dans l'image à évaluer.

Les métriques avec référence réduite Les critères de qualité avec référence réduite utilisent une description de l'image ou de la vidéo originale pour produire la note de qualité. Cette

description doit être codée et intégrée aux données de codage. Cette méthode augmente le taux de données à transmettre en laissant supposer que nous avons accès à la bande passante pour les transmettre. De plus, le décodeur devra pouvoir décoder ces segments. Pour ces raisons, nous ne les développerons pas. Le lecteur intéressé pourra se référer au chapitre 5 de l'ouvrage référencé en [Wang et Bovik, 2006].

Les métriques sans référence Les métriques sans référence sont les plus difficiles à développer dans la mesure où elles se basent sur des présomptions de ce que l'image devrait être. En effet, on tente de comparer l'image à une généralité d'images non dégradées avec le difficile choix de mettre en avant certaines caractéristiques par rapport à d'autres. Néanmoins, ces techniques sont très prisées car très peu de connaissances sur la source sont nécessaires pour pouvoir les appliquer. Ces métriques sans référence sont totalement adaptées à notre type d'application pour laquelle nous n'avons pas accès à l'image ou à la séquence originale.

4.1.3 Les analyses ascendantes (ou Bottom-Up) et descendantes (ou Top-Down)

Ces deux notions caractérisent les deux types de méthodes utilisées pour le développement de métriques objectives de qualité.

Le processus ascendant Par définition, un processus ascendant est une méthode dans laquelle des éléments de bas niveau sont définis pour permettre d'obtenir par assemblage un élément plus complexe.

Dans le cadre des métriques, l'analyse ascendante est utilisée pour modéliser chaque fonction du SVH dans le but de reproduire le complexe système visuel humain. Ces fonctions du SVH sont souvent la combinaison des caractéristiques bas niveaux. Ce type de méthode a l'avantage de se rapprocher le plus possible de la perception humaine mais requiert des temps de calcul non négligeables.

Le processus descendant Par définition, un processus descendant est le fait d'analyser un système en allant du général au particulier à condition d'avoir une bonne vision d'ensemble de la solution du problème à résoudre.

Dans le cadre des métriques, l'analyse descendante, permet de considérer la qualité en modélisant globalement ce que le SVH perçoit.

Remarque : La définition de ces deux termes est difficile car il n'existe pas de frontière nette entre les deux. D'une part, à cause de la complexité du système visuel humain, une modélisation ascendante complète de chacune des parties du SVH est impossible. Des hypothèses sont donc faites pour modéliser ce qui n'est pas tout à fait connu, et pour modéliser ce qui est difficile à intégrer de manière complète. D'autre part, pour une approche descendante, faire de bonnes hypothèses de manière à simuler correctement le système visuel humain sous-entend une bonne compréhension de chacune des caractéristiques du SVH.

4.1.4 Exemple de métriques objectives

Nous proposons d'illustrer les notions que nous avons définies avec deux exemples de métriques.

Exemple d'une métrique avec référence illustrant une approche descendante A l'inverse des critères classiques de traitement du signal ou de mesure de distance qui sont très généralistes, Wang a défini deux critères de qualité, appelées UQI (Universal Quality Index) [Wang, 2002] et SSIM (Structural SIMilarity) [Wang et al., 2004] utilisant l'information structurelle présente dans les images. L'intérêt est de ne pas reproduire les erreurs des mesures de fidélité qui comparent les informations pixel à pixel.

Dans le critère SSIM, à chaque pixel $I(x,y)$ d'une image I , on associe :

- la valeur moyenne μ_{xy} du signal sur un voisinage de taille $11*11$ pixels.
- l'écart type σ_{xy} sur ce même voisinage, indicateur de contraste
- l'information structurelle $s(x,y)$ définie par :

$$s(x, y) = \frac{I(x, y) - \mu(x, y)}{\sigma(x, y)} \quad (3.15)$$

Le calcul de l'information structurelle va permettre d'effectuer une corrélation entre les motifs de deux images, indépendamment de la luminance locale (qui est soustraite) et du contraste (qui est normalisé).

Ainsi, le calcul de similarité locale entre un échantillon p_0 de l'image originale I_0 et un échantillon p_1 de l'image dégradée I_1 réunit ces trois facteurs et s'exprime par :

$$SSIM(p_0, p_1) = \frac{(2 * \mu_{p_0} \mu_{p_1} + C1) (2 * \sigma_{p_0 p_1} + C2)}{(\mu_{p_0}^2 + \mu_{p_1}^2 + C1) (\sigma_{p_0}^2 + \sigma_{p_1}^2 + C1)} \quad (3.16)$$

avec la covariance entre les images I_0 et I_1 de taille $M*N$ pixels donnée par :

$$\sigma_{p_0 p_1}^2 = \frac{1}{N-1} \sum_{x=1}^M \sum_{y=1}^N (I_0(x, y) - \bar{I}_0) (I_1(x, y) - \bar{I}_1) \quad (3.17)$$

avec $C1 = (K_1 \cdot I_{max})^2$ et $C2 = (K_2 \cdot I_{max})^2$. I_{max} est la valeur maximale d'un pixel (255 pour un codage sur 8 bits) et $(C1, C2) = (0.01, 0.03)$ (valeurs fixées empiriquement). La similarité globale entre l'image originale I_0 et l'image dégradée I_1 est la valeur moyenne des similarités locales :

$$MSSIM(I_0, I_1) = \frac{1}{MN} \sum_{x=1}^N \sum_{y=1}^M SSIM(I_0(x, y), I_1(x, y)) \quad (3.18)$$

Pour obtenir le critère UQI, il suffit de poser $C1=C2=0$. Pour accélérer le temps de calcul, il est possible de sous-échantillonner le nombre de pixels pour lesquels on calcule le SSIM.

Dans le cadre des métriques avec références, ces métriques sont aujourd'hui largement entrées en concurrence avec le MSE et le PSNR, car nous verrons que leurs résultats correspondent plus à notre perception des défauts.

Exemple d'une métrique sans référence utilisant une approche ascendante Dans [Massidda et al., 2004], les auteurs proposent une étape de pré-traitement sur la séquence reçue pour définir de quelle façon elle est perçue. Ce schéma est représenté sur la figure 3.9.

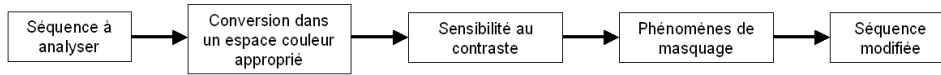


FIG. 3.9 – Exemple d’approche d’un processus ascendant : l’image est modifiée pour mettre en avant ce que le SVH perçoit.

La modélisation de chaque bloc correspond à un phénomène de bas niveaux, ce qui leur permet d’obtenir la séquence telle qu’elle est perçue. A partir de cette séquence modifiée, on peut choisir d’analyser les caractéristiques souhaitées. Les auteurs appliquent une métrique de bloc, de flou ainsi qu’une mesure des artefacts de la compensation de mouvement.

Ces modèles sont inspirés de modélisation du SVH. Ces modélisations sont complexes et difficiles à intégrer en temps réel. Le lecteur pourra se référer aux modèles utilisés par Stefan Winkler [Winkler, 1998] et Marcus Nadenau [Nadenau, 2000] sur ce sujet.

4.2 La validation des métriques

Le principe Les métriques objectives ont été conçues pour pallier aux contraintes des tests subjectifs. Néanmoins, la vérification de la validité d’une métrique objective s’obtient en analysant la corrélation entre le MOS obtenu sur une série de tests subjectifs et les notes données par la métrique sur la même série d’images ou de vidéos. Les tests subjectifs sont majoritairement des tests de mesure absolue car les résultats donnant une note moyenne par image sont plus faciles à confronter aux métriques objectives donnant également une note par image.

Le coefficient de corrélation Le coefficient de corrélation mesure le degré d’association entre deux variables X et Y . En supposant que la relation entre les variables est linéaire, que ces variables sont indépendantes et que la distribution suit une loi normale, le coefficient de corrélation de Pearson peut être calculé. Le coefficient de corrélation varie entre -1 (relation négative parfaite) et $+1$ (relation positive parfaite). Il prend la valeur de 0 s’il n’y a pas de relation linéaire entre les deux variables. /Ces protocoles et ces calculs de corrélations sont détaillés dans [VQEG, 2002].

Exemples Nous proposons de donner un exemple sur la métrique du SSIM. La figure 3.10, tirée de [Wang et al., 2004], représente pour deux bases de données d’images la corrélation obtenue entre la métrique objective SSIM et des résultats de tests subjectifs. Chaque croix ou astérisque représente une image à laquelle correspond le MOS en ordonnée et la note donnée par le SSIM en abscisse. La courbe représente une interpolation qui permet d’étudier l’existence d’une relation entre les deux variables. Plus les points sont proches de l’interpolation, plus la relation est importante. Les auteurs indiquent un coefficient de corrélation à 0.96 donc proche de 1 . A partir de cette étude, on peut donc conclure que la métrique est validée pour le type d’images utilisé dans ces tests subjectifs.

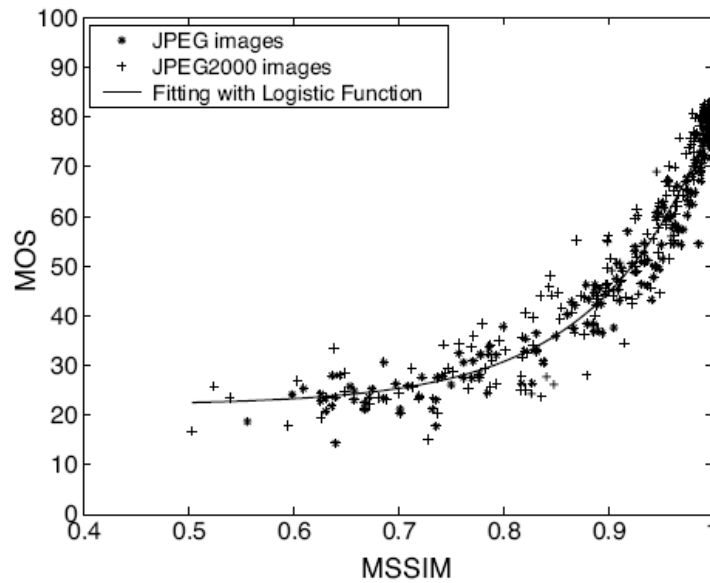


FIG. 3.10 – Figure extraite tirée de [Wang et al., 2004], qui représente la corrélation entre le MOS et le SSIM pour deux bases de données d’images.

4.3 Une mesure de qualité globale ou spécifique

De la même façon que pour les tests subjectifs, la difficulté est de savoir ce que l’on juge : la qualité de manière générale ou un artefact bien précis.

4.3.1 Une mesure de qualité globale

Mesurer l’impact des dégradations en s’aidant de la référence Des métriques telles que le SSIM sont développées pour donner un aspect global de la qualité. Ces métriques sont souvent avec référence car elles s’aident du contenu des images originales pour comparer l’impact des informations ajoutées ou supprimées sur l’image dégradée. Ces métriques peuvent être utilisées pour plusieurs artefacts mais seront moins précises dans le cas de l’étude d’un défaut spécifique.

Mesurer la dégradation la plus gênante Dans le cadre d’une métrique n’ayant pas ou peu accès à cette référence, la qualité globale est difficile à évaluer. Dans, [Quintard et Larabi, 2007], les auteurs proposent d’établir une métrique de qualité basée sur des critères perceptifs. Ils associent la notion de qualité aux notions de teinte, vivacité, contraste, saturation et naturel. Partant sur ces qualificatifs, ils mettent en place des tests subjectifs dans le but de quantifier l’impact de chacun. Ils proposent pour cela aux observateurs un questionnaire sur l’estimation de la qualité globale et sur l’estimation de chacun des qualificatifs. L’une des conclusions de ce test est la forte corrélation obtenue entre la notion de qualité et la notion de teinte. A partir de ce résultat, l’idée est donc de déterminer une métrique de qualité basée sur la teinte. Néanmoins, cela suppose que l’image n’a subi aucun bruit, aucun traitement de compression ou de redimensionnement et nécessite donc des connaissances sur la source de l’image.

Regrouper plusieurs mesures Si plusieurs défauts sont caractéristiques de la qualité d'une image, on peut envisager de regrouper les métriques propres à chaque défaut. La difficulté est alors d'associer un poids à chacune des métriques et de les regrouper suivant un type de combinaison précis. L'ajustement de ces métriques pour donner une note globale se fait souvent par une interpolation des résultats suivant une série de tests subjectifs [Wang et al., 2002], [Perra et al., 2005]. L'inconvénient de cette méthode est que la métrique devient dépendante du panel d'images utilisé pour les tests subjectifs.

Avant de s'interroger sur le regroupement de métriques, on doit être capable d'établir une métrique spécifique pour un artefact.

4.3.2 La mesure spécifique d'un artefact

Les avantages La mesure spécifique d'un artefact permet d'utiliser ses caractéristiques ce qui la rendra beaucoup plus précise qu'une métrique de qualité globale.

Les inconvénients En revanche, cette métrique ne pourra être utilisée que par rapport à la caractéristique de l'image ciblée et devra être utilisée avec précaution.

4.3.3 La mesure de la qualité par image ou par séquence

La mesure de la qualité par image Ce type de mesure permet d'avoir un contrôle de chaque image d'une séquence. Si on anticipe sur l'aspect correction des artefacts, une bonne correction commence par une bonne détection des artefacts. Ce type de mesure peut donc être exploité si on souhaite appliquer une correction image par image en fonction de la qualité retournée. En revanche, cette mesure de qualité image par image ne reflète pas totalement la qualité perçue car elle ne tient pas compte de l'aspect temporel.

La mesure de la qualité par séquence Cette technique est dédiée à étudier la qualité de la séquence globale. En intégrant l'aspect temporel, la mesure de la qualité d'une vidéo donne des résultats plus justes et plus liés à notre SVH que la mesure de la qualité image par image. Par exemple, la société Genista de Suisse a commercialisé la PQoS [Genista,] qui se présente comme un boîtier analysant les séquences vidéo et attribuant à ces séquences une note de qualité. La métrique "No Reference Objective Quality Metric" de Caviedes et al [Caviedes et Oberti, 2003] permet également d'attribuer une note par séquence.

En revanche, même si ce type de méthode donne une valeur de qualité globale, il ne pourra pas guider aussi efficacement une éventuelle correction qu'une estimation de la qualité image par image.

Le choix de ces paramètres dépend donc de l'application et de l'usage envisagé de la métrique.

Nous avons défini les différentes catégories de mesures. Le but était de comprendre le choix de la méthode la plus exploitable dans notre contexte. Nous expliquons ce choix dans la section suivante.

5 Retour au contexte

Nous avons défini un contexte de travail, des contraintes liées à ce contexte et une problématique. A partir de ces caractéristiques, nous allons définir le choix de la catégorie de métrique et la stratégie envisagée pour nous aider à résoudre la problématique de notre travail.

5.1 Choix de la catégorie de métrique

Dans notre cas d'étude, n'ayant pas accès aux images ou aux séquences d'origines, nous ciblons notre étude vers les métriques sans référence. Pour étudier la qualité des images diffusées, les métriques objectives de qualité sont plus appropriées que les métriques de fidélité. Pour une métrique, les contraintes ne sont pas aussi strictes que pour les algorithmes de correction. En effet, la contrainte de temps réel n'est pas nécessaire pour le développement d'une métrique. Par contre, ces métriques pourront servir au contrôle de la qualité et pourront être implémentées. Pour cela, nous garderons à l'esprit la contrainte de faible coût. Cette contrainte ne nous permettra pas d'intégrer des modèles psychovisuels et guidera notre étude sur des approches descendantes.

5.2 La stratégie envisagée

La problématique de la réduction des artefacts de compression MPEG-2 est la suivante : comment réduire les artefacts de compression sans faire apparaître les artefacts de correction. Pour cela, nous avons choisi de développer des métriques capables de définir d'une part la visibilité des artefacts de compression et d'autre part, la visibilité des artefacts de correction.

Les artefacts de compression Nous avons souligné dans les chapitres précédents que l'effet de bloc et l'effet de ringing étaient les deux principaux artefacts gênants dus à la compression. Dans le cadre de la compression, nous recherchons une mesure de l'effet de bloc. Pour l'effet de ringing que nous avons pourtant qualifié comme un artefact tout aussi important que l'effet de bloc, nous avons observé sur plusieurs images et séquences que la perception de cette dégradation était parallèle à la perception de l'effet de bloc. Pour cette raison et pour une question de temps, nous nous sommes restreints à l'étude de l'effet de bloc.

Les artefacts de correction Dans le chapitre précédent, nous avons insisté sur le fait qu'une image compressée avec des artefacts visibles pouvait être corrigée en atténuant les artefacts mais en en faisant apparaître d'autres et qu'il était donc difficile de juger de la qualité finale sans émettre d'hypothèses. La figure 3.11 met en évidence ce genre de problèmes. Après correction, l'image de droite n'a plus d'effet de bloc. Néanmoins, elle est de mauvaise qualité car tous les détails ont été lissés.

Le flou étant la principale conséquence de la correction de l'effet de bloc et de l'effet de ringing, nous l'avons défini comme étant l'artefact de correction le plus probable.



FIG. 3.11 – A gauche : Zoom sur une image compressée, les frontières de bloc sont visibles. A droite, zoom sur une image corrigée par filtrage : les frontières de bloc ne sont plus visibles mais l'image est floue.

La qualité image par image Notre orientation vers l'étude de l'effet de bloc et de l'effet de flou montre notre intérêt pour des artefacts de compression spatiale. Même si ces artefacts ne sont pas perçus de la même façon en spatial et en temporel, nous avons souhaité dans un premier temps être capable d'étudier leur impact image par image en gardant à l'esprit que nous pourrions les réutiliser pour la correction. Nous nous sommes donc dirigés vers une évaluation de la qualité image par image.

Dans le cadre de la compression, nous proposons un critère de contrôle de qualité basé sur deux métriques spécifiques. L'effet de bloc étant le principal artefact dû à la compression et l'effet de flou étant la principale cause d'une correction trop forte, ces paramètres semblent prépondérants dans l'estimation de la qualité d'une image compressée, corrigée ou non.

Ce chapitre introductif sur l'estimation de la qualité nous a permis de distinguer les notions de fidélité et de qualité mises en parallèle avec les notions de reconstruction et d'amélioration d'images. Le moyen le plus valide pour estimer la qualité est encore de récolter les différents jugements de représentants de la population concernés par l'application. Ces tests subjectifs sont normalisés par l'ITU et requièrent de la rigueur et du temps. Soulignant ces contraintes, nous avons développé l'autre possibilité qui consiste à utiliser des métriques objectives.

Parmi ces métriques, nous avons montré les faiblesses des métriques de fidélité. L'intégration de quelques caractéristiques du SVH permet d'obtenir la séquence telle qu'elle est perçue et est utilisée par les métriques de qualité pour améliorer les résultats. Néanmoins, l'efficacité de la modélisation du SVH utilisée est liée à sa complexité.

Puis, nous avons différencié trois familles de métriques : avec référence, avec référence réduite et sans référence. La référence apporte une aide dans l'estimation de la qualité mais restreint son domaine d'applications. Au contraire, les métriques sans référence ont un large domaine d'applications, mais comme aucune caractéristique de l'image non dégradée n'est connue, ces métriques sont souvent spécifiques à un artefact.

Enfin, nous avons également souligné le fait que ces métriques objectives, qui tendent à remplacer les tests subjectifs, sont soumises à ces mêmes tests pour être validées. L'intérêt apparaît dans le caractère répétitif et automatique de leur utilisation une fois qu'elles ont été validées.

Nous avons vu que dans le cadre de la compression, le contrôle de la qualité s'effectue principalement sur deux catégories d'artefacts :

- les artefacts de compression.*
- les artefacts de correction.*

Le contrôle de la qualité s'effectuera donc par l'application de deux métriques objectives sans référence : la métrique de l'effet de bloc pour contrôler la qualité de la compression et la métrique du flou pour contrôler la qualité des corrections. L'élaboration de ces deux métriques est l'objet de l'étude des deux chapitres suivants.

Chapitre 4

La mesure de l'effet de bloc

Dans ce chapitre, nous cherchons à estimer la gêne visuelle occasionnée par l'effet de bloc. Or, savoir localiser un artefact et estimer sa visibilité est la première étape d'une bonne correction. Pour cette raison, nous retrouverons dans ce chapitre de nombreuses similitudes avec les méthodes de correction de l'effet de bloc. De plus, nous garderons à l'esprit que le développement de cette métrique, pour estimer la qualité des images compressées, pourra servir au contrôle d'une correction.

Pour exploiter ces métriques dans notre contexte, c'est-à-dire analyser l'effet de bloc sur les images diffusées, nous avons ciblé les métriques sans référence. Nous commencerons donc par un état de l'art des métriques de l'effet de bloc sans référence pour lesquelles nous décrirons les points forts et les points faibles. En reprenant la structure de ces métriques, nous proposerons des améliorations aux limitations évoquées. Cela nous amènera à décrire notre métrique de l'effet de bloc : le BLE (Block Level Estimator). L'efficacité du BLE sera justifiée par une corrélation effectuée avec des tests subjectifs et par la comparaison des résultats obtenus avec d'autres métriques.

1 Etat de l'art des mesures de bloc sans référence

Les mesures de bloc sans référence peuvent être développées dans le domaine spatial ou dans le domaine fréquentiel. Nous allons décrire des métriques développées dans ces deux domaines et sélectionner les plus adaptées à notre contexte.

1.1 Dans le domaine fréquentiel

La FFT renseigne sur le comportement des fréquences d'une image. En considérant que les frontières de bloc sont de l'information répétée à des fréquences constantes, l'utilisation du domaine fréquentiel se retrouve dans plusieurs métriques dont celle de Wang et al [Wang et al., 2000].

Description de la méthode de Wang et al [Wang et al., 2000] La figure 4.1 schématise les différentes étapes de cette métrique. L'idée clef est de modéliser le signal contenant les effets de bloc par un signal sans effet de bloc superposé à un signal pur d'effet de bloc. Pour cela, les étapes suivantes sont effectuées :

- Création de l'image différence (en valeur absolue) ligne à ligne et de l'image différence (en valeur absolue) colonne à colonne.
- Transformation des images à deux dimensions en des images à une dimension 1D.
- Calcul de la transformée de Fourier FFT sur les signaux à une dimension 1D.
- Etude des spectres de chaque signal 1D. Par exemple, pour les frontières verticales, les frontières de blocs sont représentées sur le spectre par des pics à une fréquence correspondant au découpage de l'image en blocs 8x8. Pour modéliser le signal sans effet de bloc, un filtre médian permet de supprimer ces pics.
- Puis, la comparaison entre ce spectre de l'image avec l'effet de bloc et le spectre filtré, permet d'obtenir l'effet de bloc vertical. Le même principe est appliqué pour les frontières horizontales.
- Enfin, la pondération et la sommation de la mesure de l'effet de bloc vertical et de l'effet de bloc horizontal permettent d'obtenir l'estimation finale de l'effet de bloc.

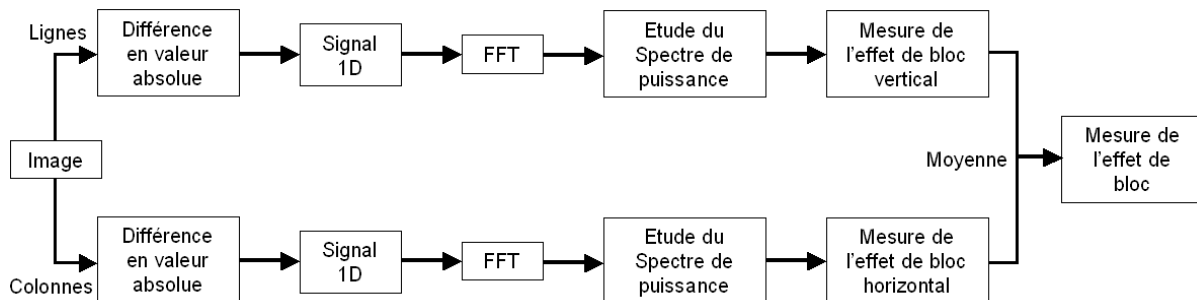


FIG. 4.1 – Principe de mesure d'effet de blocs par Wang et al [Wang et al., 2000].

Les résultats Pour une série d'images compressées à différents taux de qualité, nous allons observer le comportement de ce critère. Les tests sont effectués uniquement sur la luminance. La figure 4.2 présente les images compressées à différents taux, et la figure 4.3 représente le résultat du critère de Wang et al et le résultat fourni par le PSNR. Le facteur de qualité Q donné par *Microsoft Photo Editor* dépend uniquement du taux de compression. Or, comme la qualité est liée à la fois à la compression et au contenu de l'image, nous n'utiliserons pas cet indicateur pour évaluer la qualité mais juste comme une dénomination pour la description des graphiques.

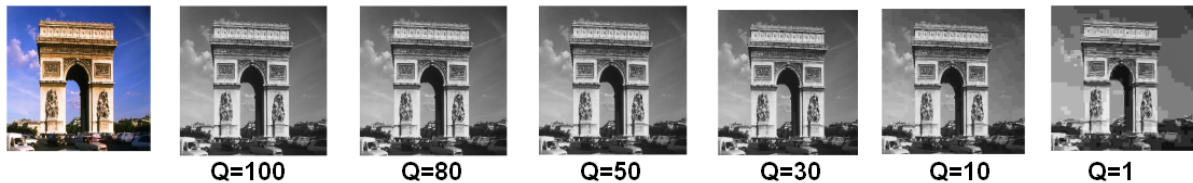


FIG. 4.2 – Images-tests que nous utilisons pour l'évaluation de la métrique de Wang et al [Wang et al., 2000].

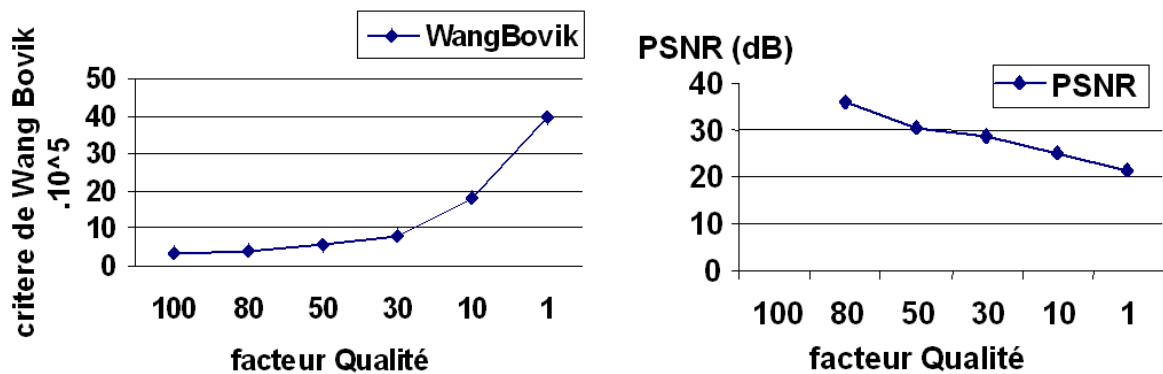


FIG. 4.3 – Résultats de la métrique de Wang et al [Wang et al., 2000] et du PSNR sur les images présentées sur la figure 4.2.

Visuellement, les effets de blocs ne se voient clairement que pour $Q=10$ et $Q=1$. Le critère de Wang Bovik augmente avec la dégradation de l'image. La courbe nous montre non seulement cette augmentation mais fait ressortir les valeurs $Q=30$, $Q=10$ et $Q=1$ pour lesquelles la dégradation semble devenir plus importante. Ces résultats sont en cohérence avec notre jugement visuel de la gêne occasionnée par l'effet de bloc. En revanche, même si le PSNR donne bien l'information de dégradation, celle-ci ne fait pas ressortir les degrés de gêne causés par l'effet de bloc.

D'autres méthodes ont également été proposées dans le domaine fréquentiel. Par exemple, dans [Gao et al., 2002], l'analyse du comportement des pixels le long des frontières s'effectue dans le domaine fréquentiel et les seuils de visibilité de ces frontières sont estimés à partir du coefficient de quantification QP . Comparée au PSNR et à des tests subjectifs, cette métrique semble efficace.

L'inconvénient du domaine fréquentiel Pour exploiter ces deux méthodes, nous serions contraints de calculer la FFT, ce qui représente une étape coûteuse en terme de calculs.

1.2 Dans le domaine spatial

Nous nous sommes plus naturellement tournées vers des métriques développées dans le domaine spatial. Ces métriques sont de plus en plus demandées pour satisfaire les conditions d'implémentation à faible coût en vue de les utiliser pour la correction de l'effet de bloc.

1.2.1 L'analyse des frontières et du contenu des blocs

Le principe est toujours centré sur l'analyse du comportement des pixels au niveau de la frontière et l'analyse du comportement des pixels à l'intérieur des blocs adjacents à la frontière. Nous allons décrire trois méthodes différentes permettant de réaliser cette étape.

Description de la méthode de Wang et al [Wang et al., 2002] Cette méthode est une alternative à la méthode [Wang et al., 2000] développée dans le domaine fréquentiel. Le principe est le suivant :

Soit $x(m, n)$ le pixel à la position m et n d'une image où :

$$m \in [1, M] \text{ et } n \in [1, N] \quad (4.1)$$

avec M , le nombre de lignes et N le nombre de colonnes de l'image.

Soit $d_v(m, n)$ la matrice représentant l'image différence suivant les colonnes telle que :

$$d_v(m, n) = x(m, n+1) - x(m, n) \quad m \in [1, M] \text{ et } n \in [1, N-1] \quad (4.2)$$

Pour analyser les frontières verticales, le principe consiste à sommer les informations au niveau des frontières verticales sur toute l'image :

$$B_v = \frac{1}{M \lfloor N/8 \rfloor - 1} \sum_{i=1}^M \sum_{j=1}^{\lfloor N/8 \rfloor - 1} |d_v(i, 8j)| \quad (4.3)$$

et au niveau du contenu des blocs :

$$A_v = \frac{1}{7} \left[\frac{8}{M(N-1)} \sum_{i=1}^M \sum_{j=1}^{N-1} |d_v(i, j)| - B_h \right] \quad (4.4)$$

Puis, une détection des passages par zéro ou "zero-crossing" est obtenue avec la formule suivante :

$$Z_v = \frac{1}{M(N-2)} \sum_{i=1}^M \sum_{j=1}^{N-2} z_h(m, n) \quad (4.5)$$

avec

$$z_h(m, n) = \begin{cases} 1 & \text{si un passage à 0 est trouvé à la position } d_h(m, n) \\ 0 & \text{sinon} \end{cases} \quad (4.6)$$

Puis, des méthodes similaires sont effectuées pour analyser les frontières horizontales et obtenir B_h, A_h, Z_h . Enfin, les valeurs correspondant aux frontières verticales et aux frontières horizontales sont moyennées pour donner :

$$A = \frac{A_v + A_h}{2} \quad B = \frac{B_v + B_h}{2} \quad (4.7)$$

Ces valeurs sont ensuite pondérées pour donner une valeur finale de la métrique. Nous reviendrons sur cette pondération dans la suite.

L'inconvénient de la méthode Wang et al L'inconvénient de cette métrique est que la gêne de l'effet de bloc n'est pas étudiée frontière par frontière mais directement sur l'ensemble de l'image. Donc, toutes les informations au niveau de la frontière sont prises en compte et cela augmente le risque de confondre les frontières de bloc avec les informations de l'image.

Description de la métrique de Pan et al [Pan et al., 2004] Cette méthode étudie la visibilité de chaque frontière entre les blocs. Cette technique permet d'être implémentée dans les algorithmes temps réel qui parcourent l'image bloc par bloc. Le principe utilisé dans [Pan et al., 2004] est représenté sur la figure 4.4 : lors de l'étude du bloc A, seuls les blocs B et C auront été parcourus. La visibilité du bloc A sera donnée par l'étude de la vision de la frontière verticale entre A et B et horizontale entre A et C, puis, les 2 autres frontières seront étudiées lors de l'analyse du bloc D et du bloc E. Leur méthode permet ainsi de parcourir une seule fois l'image.

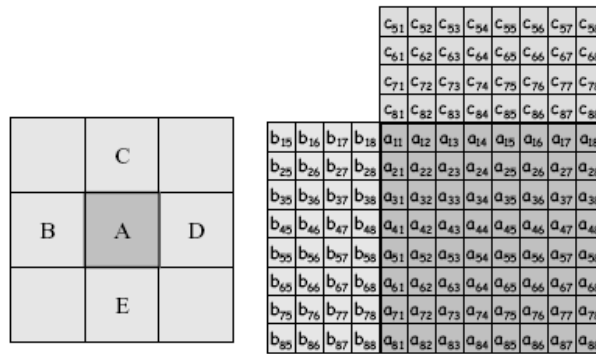


FIG. 4.4 – Représentation d'un bloc 8*8 et de ses blocs voisins. (Figure extraite de [Pan et al., 2004])

Ils définissent ensuite la visibilité de la frontière verticale entre A et B en calculant le coefficient B_h associé :

$$B_h = \begin{cases} \frac{N_h}{D_h} & \text{si } D_h \neq 0 \\ 0 & \text{si } D_h = 0 \end{cases} \quad (4.8)$$

avec N_h et D_h définis par :

$$N_h = \gamma_1 \times \sum_{i=1}^8 |a_{i1} - b_{i8}| \quad (4.9)$$

$$D_h = \gamma_2 \times \left(\sum_{i=1}^8 \sum_{j=5}^7 |b_{i(j+1)} - b_{ij}| + \sum_{j=1}^3 |a_{i(j+1)} - a_{ij}| \right) + \sum_{i=1}^8 |a_{i1} - b_{i8}| \quad (4.10)$$

Les valeurs de γ_1 et γ_2 sont trouvées de telles façons à normaliser la métrique sur une échelle de 0 à 10. Ces formules prennent en compte la valeur moyenne de la variation de l'information au niveau de la frontière par rapport à la variation d'informations du contenu des blocs adjacents à la frontière.

L'inconvénient de la méthode Pan et al L'inconvénient que nous allons présenter est également valable pour la métrique de Wang et al. Leurs analyses de la variation de l'information au niveau de la frontière et dans le contenu des blocs faussent les résultats.

En effet, ces deux méthodes ne sont pas capables de distinguer le cas a du cas b présenté sur la première étape de la figure 4.5. Pour l'étude du contenu des blocs, la variation de l'information qui se calcule ligne par ligne sera la même. En effet, la différence en valeur absolue colonne à colonne sera la même, c'est-à-dire nulle pour tous les pixels contenus dans les blocs. Cela est mis en évidence sur l'étape 2 de la figure 4.5). En revanche, au niveau de la frontière, la variation d'informations est différente.

Mais, le fait de sommer ou de moyenniser l'information dans un signal à une dimension rend identiques les 2 valeurs finales : ceci est bien schématisé sur l'étape 3 de la figure 4.5).

Au final, le cas b) sera analysé par les métriques comme le cas a) alors qu'il n'y a pas de frontière de bloc dans le cas b).

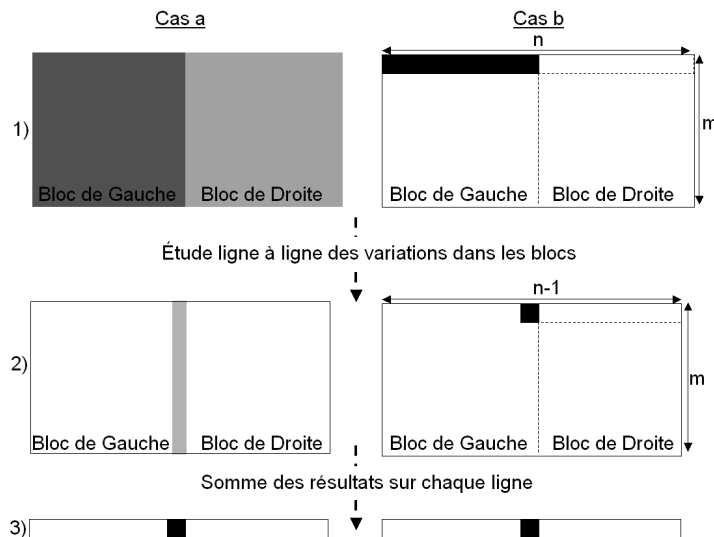


FIG. 4.5 – Possible confusion entre une frontière et un contour : 1) Représentation de deux paires de blocs de contenus différents. 2) Mise en évidence des variations d'informations pour chaque ligne de ces blocs. 3) Somme des résultats obtenus sur chaque ligne.

Description de la métrique de Perra et al [Perra et al., 2005] Nous allons voir que cette méthode est robuste à la limitation précédemment évoquée. L'étape de détection est effectuée sur les images D_x et D_y , obtenues respectivement par la convolution de l'opérateur de Sobel vertical M_x et horizontal M_y sur l'image I . Sur ces images résultantes D_x et D_y , ils considèrent des blocs de taille 8×8 sur lesquels ils définissent trois ensembles Ω_{1V} , Ω_{2V} et Ω_2 représentés sur la figure 4.6.

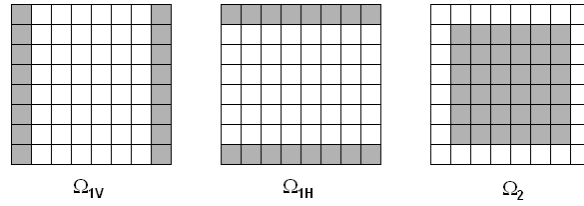


FIG. 4.6 – Les éléments en gris définissent les ensembles Ω_{1V} , Ω_{2V} et Ω_2 dans un bloc de 8*8 pixels qui a été convolué à un filtre de Sobel.

Ces sous-ensembles représentent respectivement l'analyse des frontières verticales, l'analyse des frontières horizontales et l'analyse du contenu des blocs. A partir de ces informations, la mesure de l'effet de bloc est trouvée selon les équations suivantes :

$$s_1 = \frac{1}{N_1} \left(\sum_{(i,j) \in \Omega_{1V}} \frac{|D_x(i,j)|}{\max(D_x)} + \sum_{(i,j) \in \Omega_{1H}} \frac{|D_y(i,j)|}{\max(D_y)} \right) \quad (4.11)$$

$$s_2 = \frac{1}{N_2} \sum_{(i,j) \in \Omega_2} \frac{D(i,j)}{\max(D)} \quad (4.12)$$

avec

$$D(i,j) = \sqrt{D_x(i,j)^2 + D_y(i,j)^2} \quad (4.13)$$

N_1 et N_2 sont juste deux facteurs de normalisation.

Ces formules permettent de distinguer le cas expliqué sur la figure 4.5 car en tenant compte de la valeur maximum de la variation d'information au niveau de la frontière, le cas b) sera bien distinct du cas a).

Les inconvénients Même si leur normalisation permet d'obtenir des résultats entre 0 et 1, elle ne leur permet pas de distinguer deux frontières perceptibles entre un couple de blocs n'ayant pas les mêmes niveaux d'intensité. En effet, sur l'exemple représenté sur la figure 4.7, les valeurs de s_1 et de s_2 seront les mêmes et seront égales à 1.

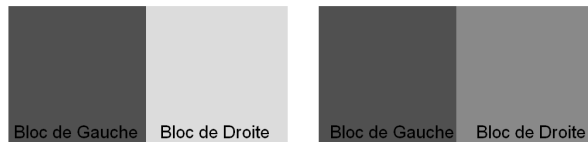


FIG. 4.7 – La frontière entre la paire de blocs de gauche est plus visible que la frontière entre la paire de blocs de droite. Pourtant, ces deux frontières sont vues de la même façon selon la normalisation de Perra et al.

Ce défaut est accentué dans la métrique proposée par [Babu et Perkis, 2005] où les auteurs définissent seulement un état binaire de visibilité d'une frontière : 0 si la frontière est invisible et 1 si la frontière est visible.

Synthèse Ces principes résument les différentes manières de mettre en avant la variation d'informations au niveau de la frontière par rapport à la variation d'informations au niveau du contenu des blocs adjacents. En revanche, ces métriques n'incluent pas de méthodes distinctes pour différencier la frontière de bloc d'un contour.

1.2.2 Différencier une frontière de bloc d'un contour

En effet, les métriques capables de différencier une frontière d'un contour ont généralement accès à la référence ou aux paramètres du décodeur. Même si cela ne fait pas partie de notre catégorie de métriques, nous allons les décrire et expliquer comment ce problème est souvent contourné par les autres métriques.

L'accès à l'image de référence Nous allons détailler le principe utilisé par la méthode [Tan et Ghanbari, 2000] qui utilise l'image de référence. Le diagramme bloc de cette méthode est représenté sur la figure 4.8). L'idée est d'effectuer une détection de contours sur l'image originale et sur l'image dégradée.

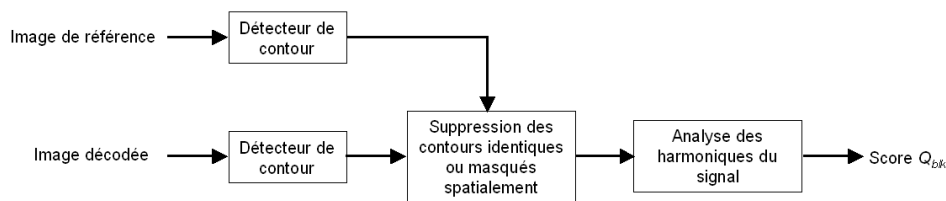


FIG. 4.8 – Diagramme bloc de la métrique de Tan et al.

Puis, la sélection des contours non communs aux deux images permet de trouver les contours restants. Ces contours restants correspondent soit aux frontières de blocs, soit aux contours modifiés par la compression. La figure 4.9 illustre l'image différence obtenue après avoir fait une estimation des contours verticaux sur l'image originale et sur l'image de référence. On note que les frontières de bloc visibles dans le ciel sont bien mises en évidence. Sur les contours restants, les auteurs tiennent compte de l'effet de masquage pour sélectionner les frontières visibles restantes.

L'accès au paramètre de quantification Connaître ce paramètre aide à différencier une frontière de bloc d'un contour. Nous avons décrit cet aspect au chapitre 2 : si le paramètre de quantification est élevé, l'information détectée est probablement une frontière de bloc. En revanche, si le paramètre de quantification est faible, l'information détectée est probablement un contour.

Dans les autres cas ? Ces deux cas ne sont pas exploitables dans notre contexte d'étude et n'ont également pas été pris en compte dans les métriques de Wang et al, Pan et al et Perra et al. En effet, cet aspect est moins gênant pour une métrique que pour une correction. La métrique de l'effet de bloc donne une note finale sur la qualité de l'image qui dépend de la valeur trouvée pour chaque frontière. Si une confusion entre une frontière de bloc et un contour se produit, la valeur sera moyennée avec les autres estimations rendant l'erreur moins importante. Néanmoins, il faut limiter ces confusions pour minimiser l'erreur.

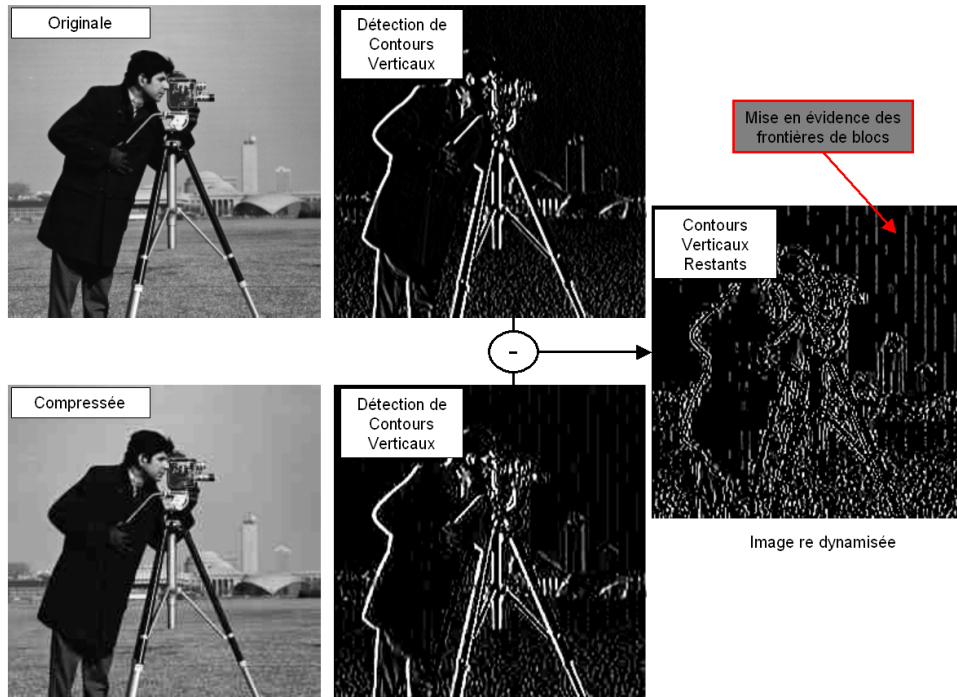


FIG. 4.9 – Exemple de mise en valeur des frontières de blocs verticales par l'application d'un filtre de Sobel et par soustraction des contours de l'image décodée aux contours de l'image originale.

1.2.3 La corrélation des métriques avec la perception

En tenant compte du comportement des pixels à l'intérieur des blocs, les métriques décrites attribuent un poids associé à la visibilité des frontières en tenant compte de l'effet de masquage. Ce principe ne donne pas des résultats aussi corrélés à la perception que les méthodes [Massidda et al., 2004] qui modélise la perception à l'aide des caractéristiques du SVH. Néanmoins, la corrélation peut être améliorée en utilisant les résultats des tests subjectifs.

L'utilisation des tests subjectifs Le principe de calibrage de la métrique en fonction des tests subjectifs est souvent utilisé. Par exemple, dans la métrique de Wang et al ([Wang et al., 2002]), les auteurs différencient les variations au niveau des frontières (variable B), au niveau des blocs (variable A), et les zones homogènes (variable Z). Puis ils calculent la valeur finale par la formule suivante :

$$S = \alpha + \beta B^{\gamma_1} A^{\gamma_2} Z^{\gamma_3} \quad (4.14)$$

Les différents paramètres α , β , γ_1 , γ_2 , et γ_3 sont estimés en fonction des résultats des tests subjectifs. Cela permet de corréler la métrique à la perception des images.

De la même façon, Perra et al [Perra et al., 2005] regroupent les informations de mesure s_1 et s_2 dans la formule suivante :

$$S = \left| \frac{s_1^\lambda - s_2^\lambda}{s_1^\lambda + s_2^\lambda} \right| \quad (4.15)$$

où le coefficient λ est ajusté en fonction des tests subjectifs de manière à obtenir une bonne corrélation.

L'inconvénient La validité des résultats obtenus va dépendre du type d'images et de vidéos utilisées pour effectuer les tests. Pour des images de contenus différents, les résultats ne seront pas aussi bien corrélés.

Nous venons de voir qu'il est difficile d'intégrer les 2 étapes (distinction entre une frontière de bloc et un contour et attribution d'un poids suivant les caractéristiques du SVH) dans une métrique sans référence n'ayant pas accès aux paramètres du décodeur et devant être développée à faible coût. Nous proposons de développer une métrique établissant un compromis entre ces différentes contraintes.

2 Le BLE : Block Level Estimator

Nous allons détailler les étapes du développement de notre métrique de bloc : le BLE. Comme pour les autres métriques, le BLE est évalué sur la composante de luminance de l'espace YCbCr.

2.1 La perception de l'effet de bloc

Pour que le BLE soit corrélé à notre perception des blocs, nous avons cherché à comprendre pourquoi le SVH était gêné par ces frontières de bloc.

2.1.1 Pourquoi perçoit-on une frontière ?

Mathématiquement parlant, l'effet de bloc résulte de la quantification des coefficients DCT et du traitement des blocs indépendamment les uns des autres. D'un point de vue perceptif, une frontière de bloc devient visible car il y a une discontinuité de l'information entre deux blocs adjacents.

On définit un bloc homogène comme étant un bloc dans lequel il n'y a aucune variation d'informations : le niveau de gris est le même pour chaque pixel du bloc. A l'inverse, on définit un bloc non-homogène comme un bloc dans lequel il existe des variations d'informations.

Sur la figure 4.10, deux paires de blocs sont représentées. La frontière est distincte dans les deux cas même si elle est plus visible pour la paire de blocs homogènes. Pour chaque paires de blocs, nous avons extrait une ligne d'informations.

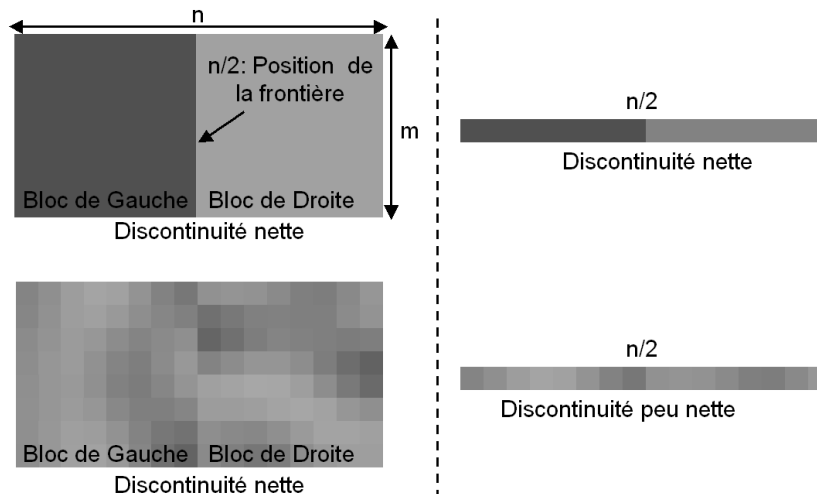


FIG. 4.10 – Les raisons de la perception d'une frontière de bloc.

Dans le cas des blocs homogènes, on remarque la discontinuité de l'information est très nette sur la ligne extraite. En revanche, pour la paire de blocs non-homogènes, cette discontinuité n'est pas aussi nette. On en déduit que la frontière est visible entre les deux blocs non-homogènes car la faible discontinuité apparaît sur chaque ligne et crée une discontinuité perceptible.

2.1.2 Comment modélise-t-on ce phénomène ?

Pour implémenter cette condition, nous proposons comme beaucoup de méthodes ([Perra et al., 2005], [Pan et al., 2004], [Wang et al., 2002]) d'étudier les variations d'informations au niveau de la frontière et à l'intérieur des deux blocs. Pour l'étude d'une frontière verticale, nous étudierons la variation d'informations colonne à colonne. Pour l'étude d'une frontière horizontale, nous étudierons la variation d'informations entre les lignes. Cela nous permettra d'utiliser un simple gradient dans une dimension et de minimiser les calculs. Nous allons détailler l'étude d'une frontière verticale pour laquelle les étapes sont représentées sur la figure 4.11.

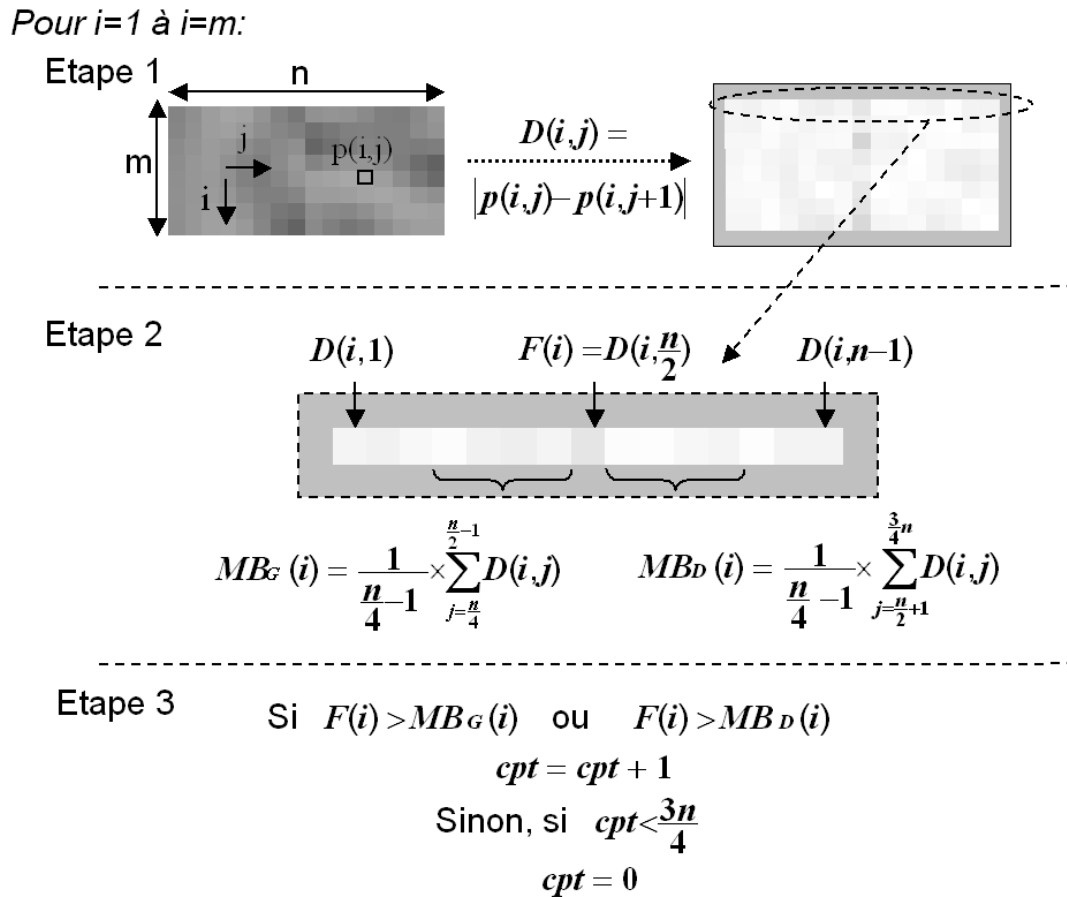


FIG. 4.11 – Les étapes de la modélisation déterminant la visibilité d'une frontière.

Etape1 Si on considère les pixels p d'un ensemble de deux blocs adjacents, la variation d'informations sera obtenue en calculant la différence en valeur absolue des pixels entre colonnes voisines. Cela est schématisé par la matrice D représentant la matrice de variation d'informations.

Etape2 Puis, pour chaque ligne de la matrice D , l'information au niveau de la frontière est comparée à l'information à l'intérieur des deux blocs. Cette étape est effectuée en calculant la moyenne de ces informations sur le bloc à gauche de la frontière MB_G et sur le bloc à droite de la frontière MB_D : c'est l'étape 2 de la figure 4.11.

Etape3 Ces valeurs sont ensuite comparées à l'information F correspondant aux variations au niveau de la frontière. En effet, si pour la ligne i , la variation d'informations au niveau de la frontière est plus importante que celle à l'intérieur des blocs, un compteur cpt est incrémenté. L'intérêt de ce compteur est double :

- il permet de vérifier que la discontinuité au niveau de la frontière est répétée.
- il permet également de s'assurer que la discontinuité est consécutive.

Remarque : de manière expérimentale, nous avons décidé que si ce compteur était supérieur à une valeur de seuil égale à au moins $3/4$ des lignes, la frontière serait définie comme perceptible. Donc si la discontinuité n'est pas obtenue sur au moins $3/4$ des lignes consécutives, le compteur est remis à 0. C'est la raison pour laquelle nous avons décidé dans l'étape 3 de remettre le compteur à 0 si les lignes n'étaient pas consécutives.

Pour résumer, grâce aux 3 étapes représentées sur la figure 4.11, nous sommes capables pour chaque frontière de conclure que :

$$\begin{cases} \text{Si } cpt > \frac{3}{4}m & \text{la frontière est considérée comme visible} \\ \text{Sinon} & \text{la frontière n'est pas visible} \end{cases} \quad (4.16)$$

L'intérêt Grâce à ce principe, nous sommes robustes à la limitation évoquée pour les métriques de Wang et Pan car nous sommes capables de distinguer le cas de la figure 4.5. En effet, pour la paire de blocs de gauche, nous trouvons que $cpt = 8$ et pour la paire de blocs de droite, nous trouvons que $cpt = 1$. Cette étape nous permet donc de définir si une frontière existe ou non.

De plus, cela permet de déterminer l'état visible ou non de chaque frontière d'une image. La figure 4.12 représente l'état binaire des frontières d'une image compressée. Une frontière blanche indique qu'une frontière perceptible existe.

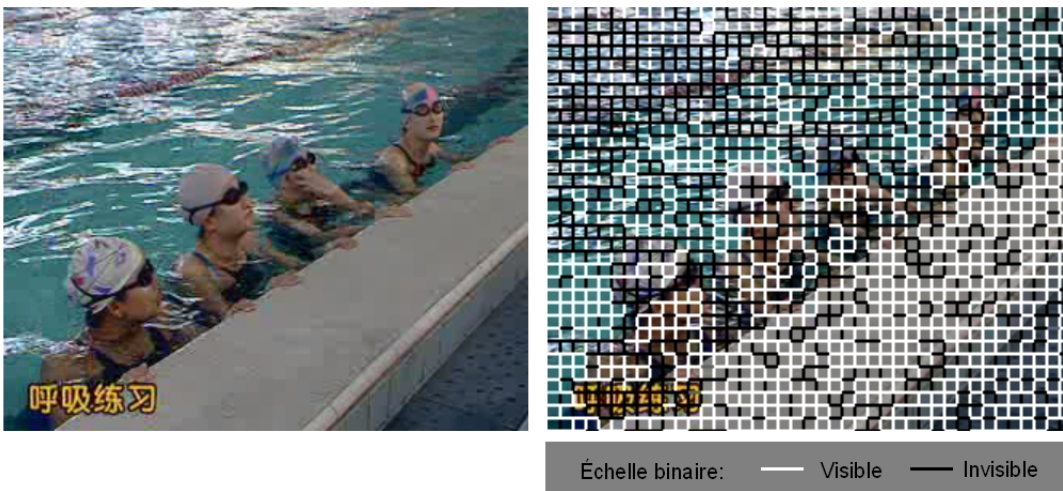


FIG. 4.12 – La représentation binaire des frontières de bloc : visible ou non-visible.

2.2 Attribuer un coefficient traduisant la gêne de l'effet de bloc

Avant de s'assurer que la frontière de bloc détectée n'est pas un contour, nous nous intéressons à l'attribution d'un poids traduisant la gêne occasionnée par l'effet de bloc. Ce poids est d'abord attribué à chacune des frontières pour ensuite donner une note de qualité finale correspondant à la gêne de l'effet de bloc occasionnée sur l'image entière.

2.2.1 Attribuer un coefficient par frontière

Si la frontière a été définie comme perceptible, il faut lui attribuer un poids suivant son degré de visibilité c'est-à-dire la gêne qu'elle occasionne.

Le principe Nous définissons comme $V_{I,J}$ la force de visibilité d'une frontière à la position (I, J) . En supposant que les frontières sont positionnées à une période de 8 pixels verticalement et horizontalement, I et J sont des multiples de 8.

Ce poids va dépendre du contenu des blocs pour tenir compte des effets de masquage. En effet, le seul critère des différences en valeur absolue au niveau de la frontière n'est pas approprié pour définir ce poids. Par exemple, sur la figure 4.13, la somme des différences en valeur absolue de l'information au niveau de la frontière est de 208 pour le premier couple de blocs et de 160 pour le second couple de blocs. Or, ces valeurs ne correspondent pas à notre perception car la frontière de la paire de blocs de droite est plus visible que celle de la paire de blocs de gauche.

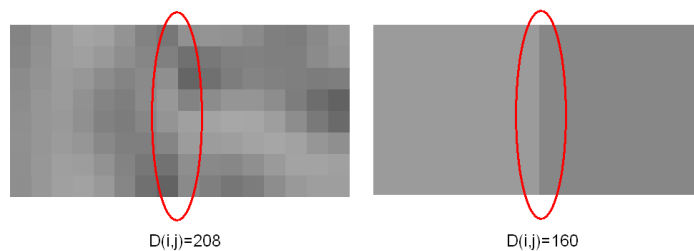


FIG. 4.13 – L'intérêt de prendre en compte le phénomène de masquage : l'information uniquement au niveau de la frontière n'est pas corrélée à notre perception.

C'est la raison pour laquelle il faut tenir compte de l'effet de masquage pour moduler la différence absolue au niveau de la frontière. Nous utilisons la moyenne des variations d'informations à l'intérieur des deux blocs mais aussi un estimateur de la variance de ces informations. Comme l'estimation de la moyenne est déjà effectuée pour estimer la visibilité de la frontière, nous proposons d'inclure le calcul de l'estimation d'une pseudo variance VB_G et VB_D au niveau de l'étape 2. La nouvelle étape 2 est représentée sur la figure 4.14.

Remarque : Notons que ce calcul sera inutile si la frontière n'est pas perceptible mais en terme de temps réel pour un algorithme, il vaut mieux effectuer un calcul inutile plutôt que de re-parcourir plusieurs fois la même information.

Ainsi, dans l'étape 3 de comparaison, nous pouvons inclure le calcul de la visibilité $V(i)$ de la frontière au niveau de la ligne i selon que l'information F est supérieure à l'un ou les deux contenus des blocs adjacents. Ce calcul de $V(i)$ contient deux étapes :

- la soustraction à F de la moyenne des variations dans les blocs voisins : cela permet de tenir compte de l'effet de masquage en luminance.
- la division du résultat par le ou les coefficients $VB_G(i)$ et $VB_D(i)$ correspondant à la ligne i : cela permet de tenir compte de l'effet de masquage dû aux variations dans le bloc.

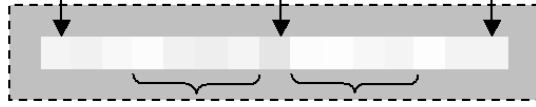
Tant que la discontinuité se répète sur des lignes consécutives, nous sommes ces valeurs $V(i)$ dans la variable $V_{I,J}$ correspondant à la visibilité globale de la frontière.

Pour $i=1$ à $i=m$:

Etape 1

Etape 2

$$D(i,1) \quad F(i) = D(i, \frac{n}{2}) \quad D(i, n-1)$$



$$MB_G(i) = \frac{1}{\frac{n}{4}-1} \times \sum_{j=\frac{n}{4}}^{\frac{n}{2}-1} D(i,j) \quad MB_D(i) = \frac{1}{\frac{n}{4}-1} \times \sum_{j=\frac{n}{2}+1}^{\frac{3n}{4}} D(i,j)$$

$$VB_G(i) = \frac{1}{\frac{n}{4}-1} \sum_{j=\frac{n}{4}}^{\frac{n}{2}-1} |MB_G(i) - D(i,j)| \quad VB_D(i) = \frac{1}{\frac{n}{4}-1} \sum_{j=\frac{n}{2}+1}^{\frac{3n}{4}} |MB_D(i) - D(i,j)|$$

Etape 3

$$\left. \begin{aligned} \text{Si } F(i) > MB_G(i) \text{ et } F(i) > MB_D(i) : V(i) &= \frac{F(i) - (MB_G(i) + MB_D(i))/2}{VB_G(i) + VB_D(i)} \\ \text{Si } F(i) > MB_G(i) : V(i) &= \frac{F(i) - MB_G(i)}{VB_G(i)} \\ \text{Si } F(i) > MB_D(i) : V(i) &= \frac{F(i) - MB_D(i)}{VB_D(i)} \end{aligned} \right\} \begin{aligned} cpt &= cpt + 1 \\ V_{I,J} &= V_{I,J} + V(i) \end{aligned}$$

$$\left. \begin{aligned} \text{Sinon, si } cpt < \frac{3n}{4} : \end{aligned} \right\} \begin{aligned} cpt &= 0 \\ V_{I,J} &= 0 \end{aligned}$$

FIG. 4.14 – Les étapes de la modélisation déterminant le coefficient représentatif de la gêne d'une frontière.

L'intérêt Cette méthode nous permet d'attribuer un poids à chaque frontière selon qu'elle est plus ou moins visible. Sur la figure 4.15, nous avons représenté les frontières suivant leur gêne. Cela nous permet par exemple de déterminer les zones de l'image où les frontières sont les moins visibles.

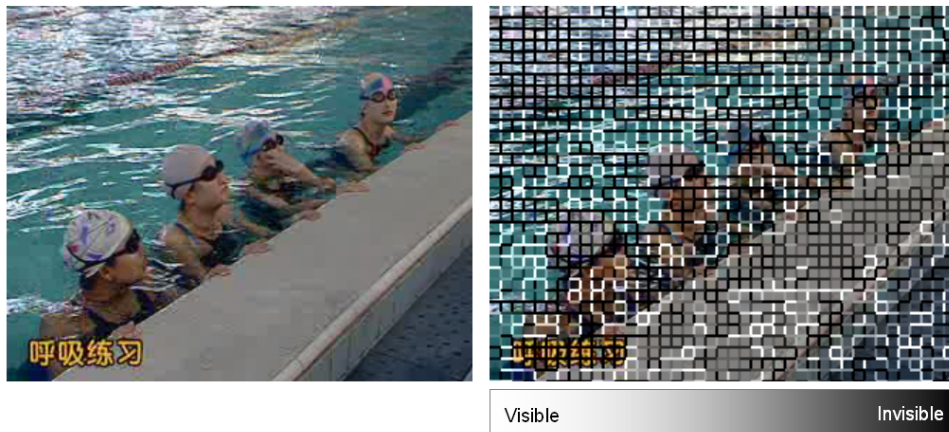


FIG. 4.15 – La représentation des différents degrés de visibilité d'une frontière.

2.2.2 Attribuer un coefficient pour l'image

Puis, pour chaque frontière $V_{I,J}$ étudiée, un compteur de frontières cpt_F est incrémenté.

$$BLE = \frac{\sum_{I,J} V_{I,J}}{cpt_F} \quad (4.17)$$

Remarque : ce compteur cpt_F est utile car la normalisation n'est pas uniquement effectuée sur l'ensemble des possibles frontières de l'image mais par rapport à un type de frontières bien définies. Nous décrivons ceci dans la suite.

2.3 Différencier l'effet de bloc du contenu de l'image

N'ayant pas accès à l'image de référence, nous proposons d'une part, une méthode pour différencier le mieux possible une frontière d'un contour et d'autre part, une analyse du contenu de l'image permettant de cibler les zones qui attirent d'avantage notre attention c'est-à-dire dans lesquelles nous serons plus sensibles à l'effet de bloc.

2.3.1 Différencier un contour d'une frontière de bloc

Nous avons vu que la distinction entre frontière de bloc et contour était envisageable si l'image sans effet de bloc, ou le paramètre de quantification, était connu. Dans notre cas, n'ayant accès à aucun de ces paramètres et ne pouvant être sûr de cette information, nous éliminons tous les cas où la frontière est probablement un contour. Pour ce faire, nous définissons la visibilité maximum d'une frontière de bloc causée par la compression. Cette valeur est déterminée de manière visuelle en recherchant statistiquement le maximum de la différence absolue obtenue au niveau de la frontière comme l'illustre la figure 4.16.



FIG. 4.16 – A gauche : composante d'intensité de l'image "vitrail". A droite : composante d'intensité de l'image "vitrail" fortement compressée. L'écart maximum trouvé entre les deux niveaux de gris de la frontière est de 32.

Ce test, effectué sur 5 images de contenus différents, nous permet de définir ce maximum à 32 niveaux de gris d'écart entre deux pixels de part et d'autre de la frontière. Nous définissons alors que si la valeur de la frontière est supérieure à $m * 32$ (m étant le nombre de lignes), il y a de grands risques que ce soit un contour. Ainsi, si :

$$\text{Si } F = \sum_{i=0}^{i=m} F(i) > \text{Seuil}_F \text{ et } \text{Seuil}_F = m * 32 \Rightarrow \text{Pas de correction.} \quad (4.18)$$

En effet, dans ce cas, nous considérons que l'information contenue dans F est un contour et nous n'en tenons pas compte.

Remarque1 : Ce test ne peut se faire que de manière visuelle pour pouvoir distinguer les frontières des contours. C'est la raison pour laquelle il n'a été effectué que sur 5 images. Néanmoins, pratiquer ces tests visuels sur un plus large panel d'images permettrait d'affiner la valeur de ce seuil.

Remarque2 : Cette méthode ne différencie pas toutes les frontières de blocs des contours mais permet d'en confondre le moins possible. Le BLE étant une métrique d'estimation de la perception de l'effet de bloc sur une image entière, les quelques confusions restantes seront moyennées avec les autres valeurs exactes $V_{I,J}$ réduisant ainsi l'erreur globale.

2.3.2 La gestion des zones homogènes

Nous cherchons, dans la mesure du possible, à tenir compte de la focalisation de notre attention sur une image. En effet, nous avons remarqué que sur une image de type texte sur un fond homogène, notre attention était principalement attirée par le texte. Or, si cette image est compressée, la zone homogène ne sera pas modifiée, seul le texte qui représente une petite partie de l'image sera compressé. Pourtant, l'effet de bloc sur cette image nous paraîtra important. Ce phénomène est représenté sur la figure 4.17.



FIG. 4.17 – La compression a des répercussions visuelles que sur les zones contenant des variations d'informations.

Pour en tenir compte dans l'estimation de l'effet de bloc, nous étudions uniquement les zones avec des variations d'informations. S'il n'y a aucune variation d'informations et que nous ne sommes pas dans le cas d'un bloc étendu (cas expliqué dans la suite), nous ne tenons pas compte de la paire de blocs étudiée. Pour ce faire, nous n'attribuons pas de valeur de BLE et nous n'incrémentons pas le compteur cpt_F associé au nombre de frontières traitées.

2.4 Cas des forts taux de compression

2.4.1 La définition du bloc étendu

Dans le cas des fortes compressions, nous sommes confrontés au phénomène de *bloc étendu*. Sur la figure 4.18, on voit que l'augmentation du niveau de compression a fait diminuer le nombre de frontières perceptibles : si on moyenne par rapport au nombre de frontières potentielles, la moyenne diminuera avec l'augmentation du taux de compression même si les frontières détectées ont un poids plus fort.



FIG. 4.18 – Le phénomène des blocs étendus.

Ce phénomène est très peu mentionné dans les métriques existantes. Or, pour une métrique qui associe un poids à chaque frontière et qui se base sur ces poids pour donner la valeur finale, l'attribution d'un poids nul à ces frontières "invisibles" fera diminuer la valeur de la métrique pour les fortes compressions, ce qui n'est pas corrélé avec notre perception de l'effet de bloc.

2.4.2 Les méthodes tenant compte du phénomène des blocs étendus

Certains auteurs ont également remarqué ce phénomène et ont cherché à le détecter. Dans [Pan et al., 2004], le principe consiste à chercher les zones où les frontières sont invisibles. La contrainte de cette méthode est qu'elle n'est pas capable de distinguer les frontières invisibles des zones réellement homogènes. Ainsi, sur une image non compressée telle que celle représentée sur la figure 4.19, leur métrique donnera une mesure de bloc élevée alors qu'aucun bloc n'existe. Dans [Gao et al., 2002], ils utilisent le même principe que pour la correction de l'effet de bloc, c'est-à-dire en parcourant plusieurs fois les mêmes blocs pour déterminer si ils sont dans une zone de compression. Or, nous avons vu que cette condition ne correspondait pas à la contrainte de faible coût.



FIG. 4.19 – Image type où la confusion est possible entre les zones homogènes de l'image et des blocs étendus.

2.4.3 Proposition d'une méthode de détection des blocs étendus

Définition Pour détecter ces frontières invisibles, nous proposons de définir un bloc étendu et pour ne pas le confondre avec une zone initialement homogène de l'image, nous le cherchons uniquement dans les zones fortement compressées. Ces zones fortement compressées sont détectables par une frontière visible entre deux blocs homogènes. Nous définissons un bloc étendu comme étant un bloc contenant une suite de frontières invisibles qui suivent une frontière visible entre deux blocs homogènes. La figure 4.20 illustre cette définition où $V_{I,J}$ est la première frontière visible entre deux blocs homogènes et $I_{I,J,k}$ sont les frontières invisibles .

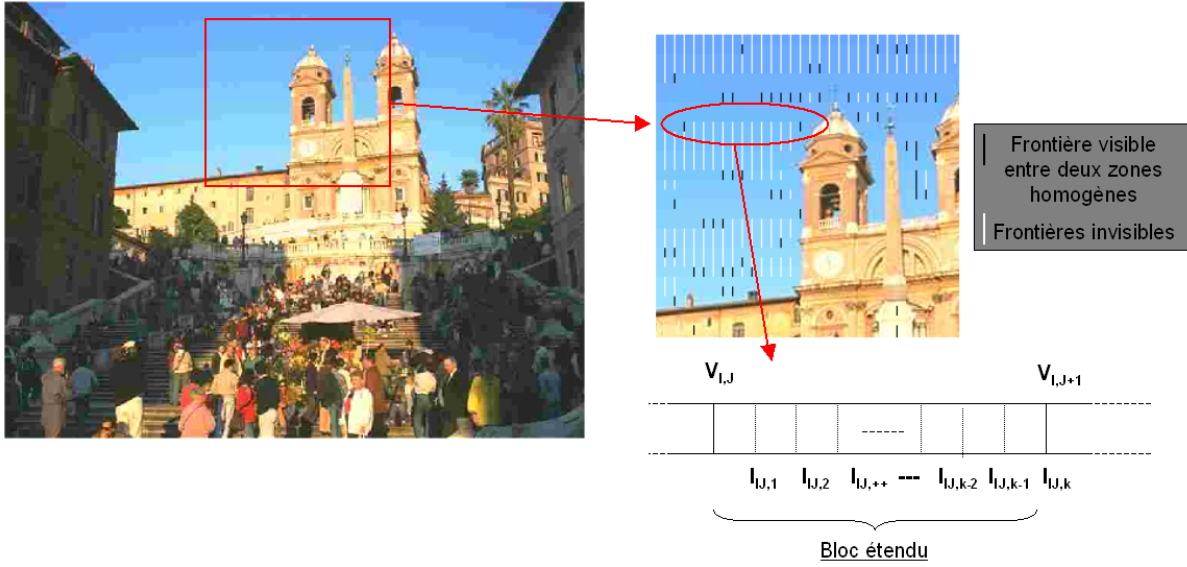


FIG. 4.20 – Définition et schématisation d'un bloc étendu.

1cm

Détection Pour la détecter algorithmiquement, deux caractéristiques sont à prendre en compte :

- la frontière de bloc $V_{I,J}$ est une frontière perceptible entre deux blocs homogènes.
- la frontière de bloc $V_{I,J+1}$ est une frontière imperceptible qui est considérée comme le début d'un bloc étendu.

Pour la première condition :

$$\begin{aligned}
 V_{I,J} &= F_{visible_homognes}(I, J) \\
 &= \begin{cases} 1 & \text{si } 0 < V_{I,J} < Seuil_F \\ & \text{et } \sum_{i=0}^m MB_G(i) < Seuil_{Hom} \\ & \text{et } \sum_{i=0}^m MB_D(i) < Seuil_{Hom} \\ 0 & \text{sinon} \end{cases} \quad (4.19)
 \end{aligned}$$

Pour la seconde condition :

$$\begin{aligned}
 V_{I,J} &= F_{invisible_homognes}(I, J) \\
 &= \begin{cases} 1 & \text{si } V_{I,J} = 0 \\ & \text{et } \sum_{i=0}^m MB_G(i) < Seuil_{Hom} \\ & \text{et } \sum_{i=0}^m MB_D(i) < Seuil_{Hom} \\ 0 & \text{sinon} \end{cases} \quad (4.20)
 \end{aligned}$$

Enfin, pour qu'une frontière de bloc appartienne à un bloc étendu :

$$V_{I,J} = I(IJ, k) = \begin{cases} 1 & \text{si } F_{invisible_homognes}(I, J) = 1 \\ & \text{et } F_{visible_homognes}(I, J) = 1 \\ 0 & \text{sinon} \end{cases}$$

La valeur du $Seuil_{Hom}$ n'est pas nulle pour tenir compte d'un éventuel bruit mais suffisamment faible pour ne pas tenir compte des variations d'informations. Nous fixons la valeur de ce seuil en fonction du nombre de lignes sommées acceptant ainsi un niveau de gris de variation d'un pixel à l'autre :

$$Seuil_{Hom} = m \quad (4.21)$$

Intérêt Sur la figure 4.21, nous avons représenté sur deux images compressées à des taux différents (couple d'images A), l'état binaire des frontières perceptibles ou non (couple d'images B et C). Le couple d'images (B) ne tient pas compte du phénomène des blocs étendus et on remarque que le nombre de frontières pour l'image la moins compressée est supérieure. En revanche, pour le couple d'images(C), on tient compte du phénomène de bloc étendu et le nombre de frontières est sensiblement équivalent pour les deux images.

Attribution d'un poids aux frontières invisibles Une fois détectée, nous associons un poids à cette ou ces frontières invisibles $I_{i,k}$ égal au poids de la première frontière visible B_n qui initialise le bloc étendu :

$$V(I_{IJ,1}) = V(I_{IJ,1}) = \dots V(I_{IJ,k}) \quad (4.22)$$

Cela permet de prendre en compte la force de visibilité des frontières. Le couple d'images (D) de la figure 4.21 illustre cette étape et montre que les frontières de l'image la plus compressée sont les plus visibles.

Remarque : Nous détectons le début plutôt que la fin d'un bloc étendu, car le parcours de l'image se fait de gauche à droite et cela nous évite ainsi de parcourir plusieurs fois le même bloc.

2.4.4 La technique d'implémentation de l'intégration des blocs étendus

De part sa définition simple, le bloc étendu est facile à repérer même dans un algorithme à faible coût et temps réel. En effet, pour éviter de revenir sur le contenu de blocs précédent (et donc d'augmenter la mémoire nécessaire), nous utilisons un simple booléen F_{ZH} tel que :

$$F_{ZH} = \begin{cases} 1 & \text{si } F_{visible_homogones}(I, J) = 1 \\ 0 & \text{sinon} \end{cases} \quad (4.23)$$

Ainsi, si la frontière suivante est détectée comme invisible, il suffira de consulter le booléen pour décider si on est dans le cas d'un bloc étendu.

2.5 L'échelle du BLE

Le BLE est donc une métrique de bloc sans référence permettant d'estimer la visibilité des frontières de bloc au sein d'une image. Les explications données pour les frontières verticales sont valables pour les frontières horizontales. Chaque frontière, verticale et horizontale, obtient un poids compris entre 0 et $32 * n$ soit 256 dans le cas fréquent des frontières sur 8 pixels. Puis tous les $V_{I,J}$ sont moyennés en fonction du nombre de frontières traitées. Le BLE est donc une métrique donnant une valeur comprise entre 0 et 256 représentant respectivement la bonne et la mauvaise qualité en terme d'effet de bloc. Néanmoins, comme le pire cas, à savoir la totalité des $V_{I,J}$ prenant la valeur de 256 est improbable, les valeurs seront plus concentrées dans la première moitié de cette échelle.

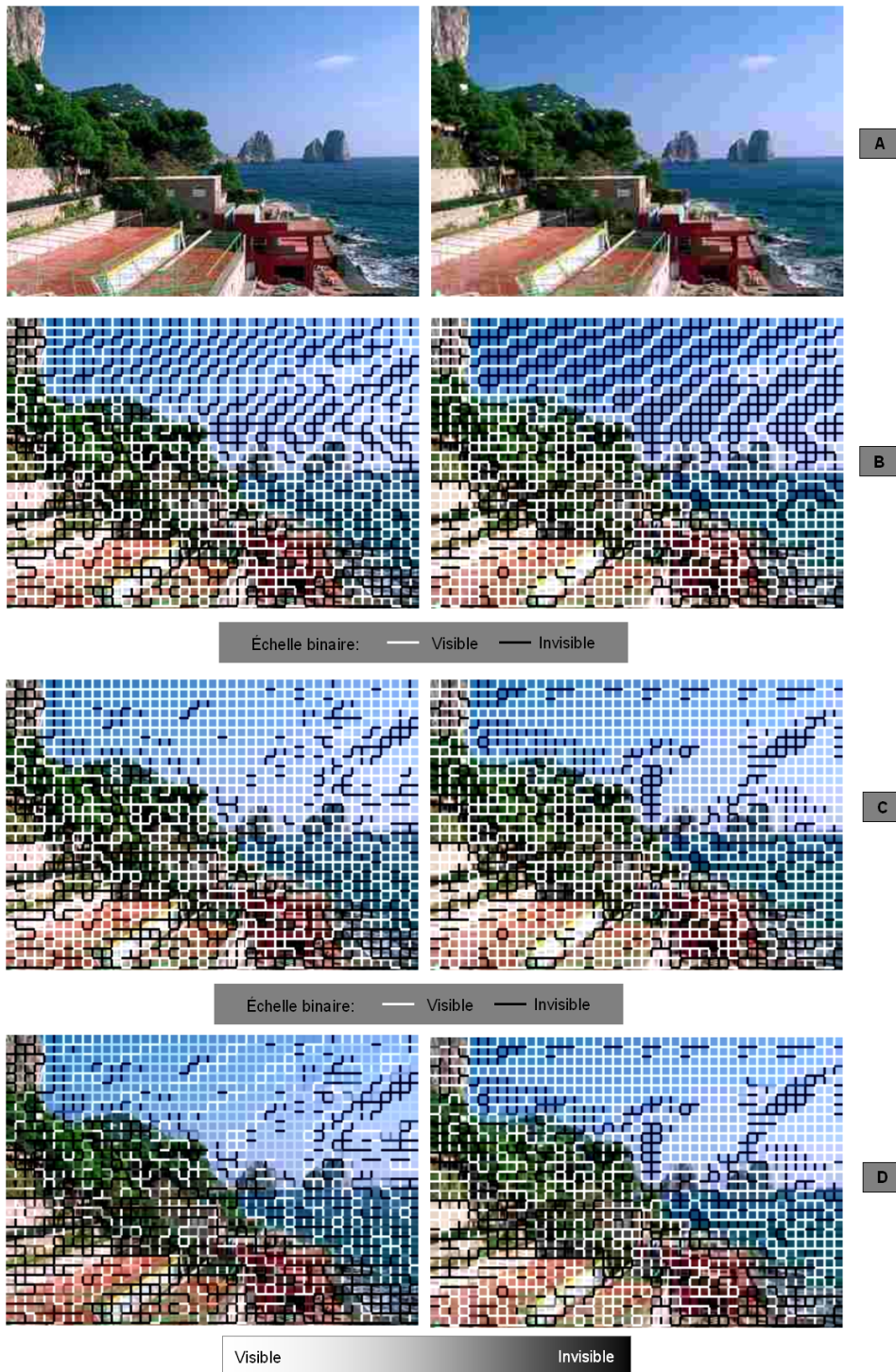


FIG. 4.21 – A) Image I1 moyennement compressée, Image I2 fortement compressée; B) Représentation de l'état binaire des frontières sans tenir compte des blocs étendus; C) Représentation de l'état binaire des frontières en tenant compte des blocs étendus; D) Représentation du degré de visibilité des frontières en tenant compte des blocs étendus;

Nous venons de présenter une métrique de bloc sans référence : le BLE. Cette métrique est basée sur le principe des méthodes existantes et intègre des améliorations au niveau de leurs limitations. Nous allons nous intéresser à sa validation.

3 La validation du BLE

Pour valider l'efficacité de ce BLE, nous avons comparé ses résultats à des tests subjectifs et aux résultats d'autres métriques de l'effet de bloc. Cette validation a été faite principalement sur des images compressées n'ayant pas subi de correction.

3.1 Les tests subjectifs

Les tests subjectifs ont été réalisés dans des conditions contrôlées mais pas forcément correspondant point par point aux conditions de l'ITU pour les raisons que nous allons expliquer.

Le matériel Faute de moyen, la salle utilisée était une salle noire sans fenêtre dans laquelle nous avons installé un écran LCD. Volontairement, nous n'avons pas choisi un écran CRT car l'effet de bloc est moins facilement perceptible sur ce type d'écran qui a tendance à filtrer les hautes fréquences.

Les images De manière à isoler l'effet de bloc et à ne pas complexifier l'information de l'image, nous avons utilisé des images compressées à différents taux en JPEG. Ces images représentent des scènes naturelles plus ou moins texturées présentées en couleur. La figure 4.22 représente un échantillon de ces images. Même si le BLE a été uniquement estimé sur la composante de luminance, nous avons choisi d'afficher les images en couleur pour le confort de l'observateur. Le test est constitué de 2 parties dans lesquelles la même série de 78 images est présentée aux observateurs. Ce test est un SSCQS (test à simple stimulus) et l'observateur note la qualité de l'image de 1 (très gênant) à 5 (imperceptible) en fonction du degré de visibilité de l'effet de bloc. Une série de 12 images est montrée en début de tests de manière à calibrer le SVH sur les niveaux de dégradations. Ces résultats ne sont pas pris en compte dans le calcul du MOS. Les images sont présentées pendant une durée de 4 secondes, puis, l'observateur a tout le temps d'attribuer une note.



FIG. 4.22 – Echantillon des images présentées dans le test subjectif.

Les observateurs Les observateurs, au nombre de 16, étaient des hommes et des femmes travaillant à STMicroelectronics. Nous avons sélectionné des personnes ayant une connaissance de l'effet de bloc pour qu'elles soient capables de le distinguer des autres artefacts. En effet, il était demandé aux observateurs de noter uniquement l'effet de bloc et non la qualité générale de l'image.

Le MOS Le Mean Opinion Score est ensuite calculé en suivant les recommandations de l'ITU.

Remarque : nous avons demandé un jugement de qualité portant uniquement sur l'effet de bloc et non sur la qualité générale des images. Ce choix a été fait pour éviter que les résultats soient biaisés par des fausses couleurs ou par l'effet de ringing. Le terme MOS sera donc ici abusivement utilisé car il ne représentera pas la qualité générale mais la qualité uniquement par rapport à l'effet de bloc.

3.2 La corrélation entre le MOS et les tests subjectifs

La figure 4.23 représente en abscisse les notes objectives obtenues par le BLE sur les 78 images et en ordonnée le MOS obtenu. Nous avons cherché la meilleure corrélation et l'avons obtenue avec un polynôme de degré 4 pour lequel le coefficient de détermination calculé avec *Excel* donne une valeur de 0.90 proche de 1.

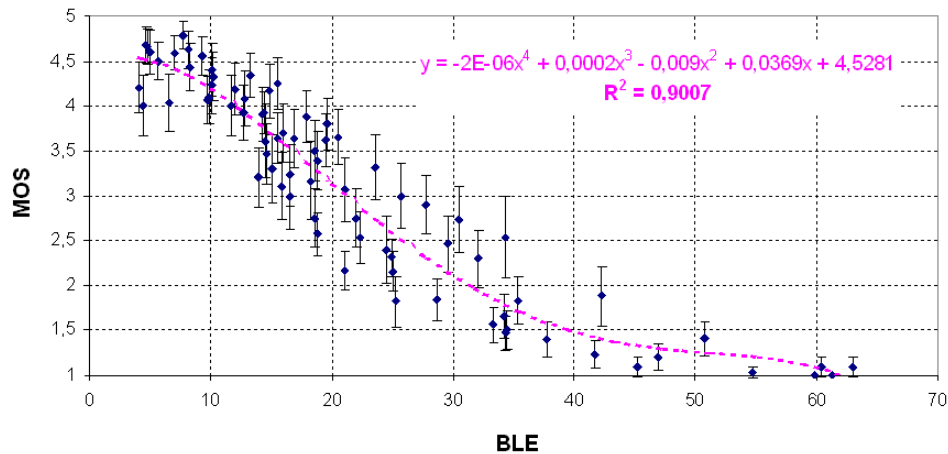


FIG. 4.23 – Corrélation entre le MOS et le BLE.

Cette corrélation nous permet donc d'associer un degré de visibilité aux valeurs de la métrique (figure 4.24) dans le cadre d'images de scènes naturelles compressées en JPEG. Conscients que cette corrélation est propre à la série d'images utilisées, nous n'avons pas normalisé la métrique par rapport à ces données.

Dégradation	Note subjective	Métrique de Flou
Imperceptible	5	5
Perceptible mais non gênant	4	10
Légèrement gênant	3	20
Gênant	2	40
Très gênant	1	>60

FIG. 4.24 – Niveaux de dégradation associés aux valeurs de la métrique.

3.3 La confrontation du BLE avec les autres métriques

3.3.1 Le BLE vs les métriques sans référence

Nous avons choisi deux métriques qui répondent aux mêmes contraintes que le BLE. Ces métriques sont la métrique M1 pour la mesure de Wang et al [Wang et al., 2002] totalement consacrée aux images JPEG et la métrique M2, pour la mesure de Perra et al [Perra et al., 2005]. La métrique M1 donne des valeurs de 1 à 10 représentant respectivement la mauvaise et la bonne qualité en terme d'effet de bloc. La métrique M2 va de 0 à 1 représentant respectivement la bonne et la mauvaise qualité en terme d'effet de blocs. Néanmoins, de part la normalisation, les valeurs sont concentrées entre 0.1 et 0.6 de manière générale. Les figures 4.25 et 4.26 montrent les corrélations obtenues avec ces métriques. Nous remarquons que la métrique M1 totalement consacrée aux images JPEG est très bien corrélée avec les tests subjectifs. En revanche, la métrique M2 est faiblement corrélée et cela peut s'expliquer par la limitation liée à la normalisation.

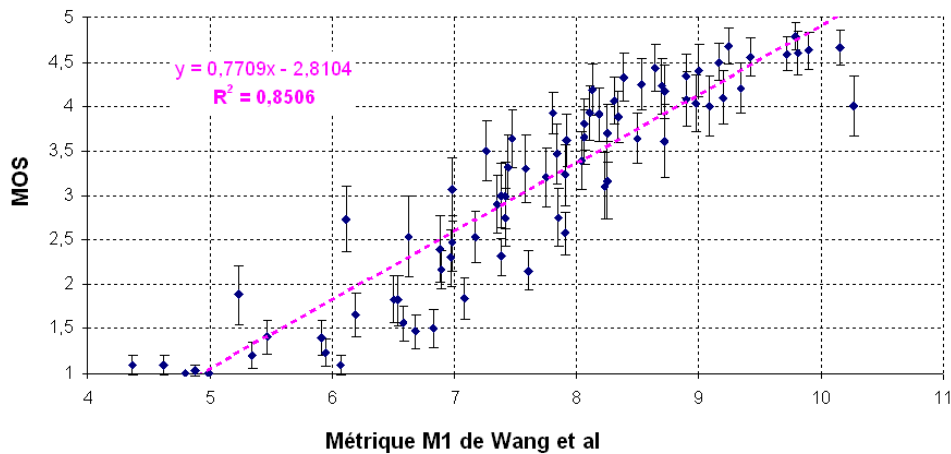


FIG. 4.25 – La métrique de Wang et al M1 vs le MOS.

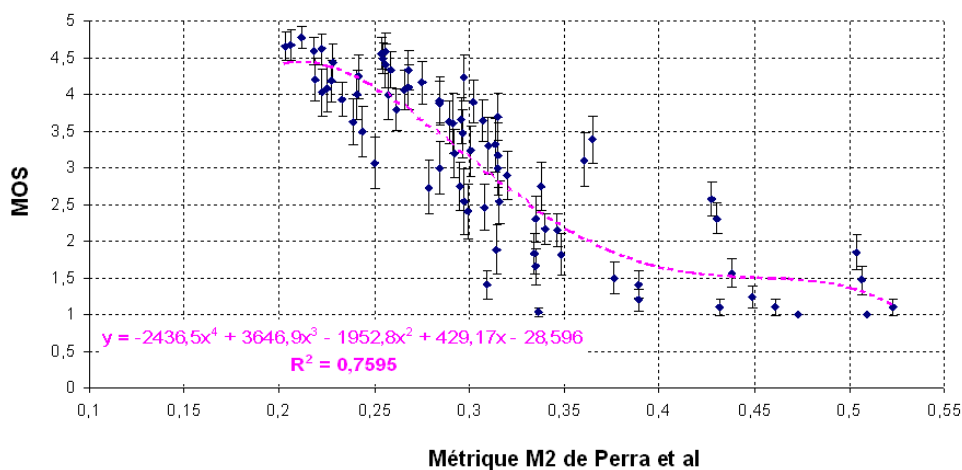


FIG. 4.26 – La métrique de Perra et al M2 vs le MOS.

Les résultats indiquent que notre métrique, avec un coefficient de corrélation égal à 0.90 contre 0.85 pour M1 et 0.78 pour M2, semble la plus corrélée à notre perception par rapport aux différentes courbes d'interpolation utilisées.

TAB. 4.1 – BLE, M1 et M2 sur les images de la figure 4.17

Métrique	Image Originale (gauche)	Image Compressée (droite)
BLE	5	26
M1	8.16	6.19
M2	0.59	0.61

TAB. 4.2 – BLE, M1 et M2 sur les images de la figure 4.18

Métrique	Image Compressée (gauche)	Image Compressée (droite)
BLE	31	60
M1	6.67	5.19
M2	0.44	0.44

3.3.2 L'efficacité de la métrique

Après avoir mis en avant l'efficacité globale de la métrique, nous allons montrer l'intérêt d'avoir tenu compte de tous nos paramètres en se référant aux limitations évoquées pour les autres métriques.

La gestion des zones homogènes Si on applique le BLE, M1 et M2 sur les deux images de la figure 4.17, nous obtenons les résultats indiqués dans le tableau 4.1.

En tenant compte que notre attention n'est pas attirée par les zones homogènes, le BLE est capable de dire que l'image de droite est très compressée et qu'il n'y a pas d'effet de bloc sur l'image de gauche. Pour la métrique M1, le phénomène est également pris en compte grâce à l'analyse du passage par zéro. Pour la métrique M2, les résultats obtenus sont proches pour les deux images et correspondent à une perception élevée de l'effet de bloc. Ceci est la conséquence de ne pas avoir de seuil de détection entre frontières et contours et de ne pas tenir compte des blocs étendus.

La prise en compte du phénomène des blocs étendus Nous avons appliqué le même principe pour le phénomène des blocs étendus en soumettant les deux images I1 et I2 de la figure 4.18 aux métriques BLE, M1 et M2.

On remarque que le BLE détecte bien la différence de visibilité de l'effet de bloc d'une image à l'autre. Pour la métrique M2, les résultats obtenus entre les deux images sont similaires. Cela est la conséquence d'une part de la normalisation de la valeur au niveau de la frontière qui rend impossible la distinction de la force des frontières entre deux zones homogènes d'une image à l'autre et d'autre part de la non-prise en compte du phénomène des blocs étendus. La métrique M1 note la différence car les poids attribués aux frontières visibles sont plus forts sur l'image de droite et viennent compenser le nombre de frontières en moins. En revanche la différence obtenue entre les 2 valeurs n'est pas représentative de la différence visuelle de perception de l'effet de bloc.

Les tests subjectifs et les comparaisons avec les autres métriques nous permettent de valider le BLE comme une mesure de l'effet de bloc sans référence corrélée à notre perception. Nous allons maintenant lister ses limitations.

4 Les limitations du BLE

Deux principales limitations retrouvées dans beaucoup de métriques et dans la nôtre sont présentées et justifiées dans cette section.

4.1 La position des frontières

La position des frontières pourra être modifiée dans le cadre d'une compression de type AVC MPEG-4 ou suite à un changement de dimension de l'image.

4.1.1 Les nouveaux codecs

Même si notre métrique est qualifiée de métrique sans référence qui n'utilise aucun paramètre du décodeur, nous faisons une hypothèse majeure sur la position des frontières de bloc. Travaillant essentiellement sur du MPEG-2, nous avons supposé que les frontières de bloc étaient toujours positionnées sur une grille de 8 pixels par 8 pixels. Or cette hypothèse n'est plus vraie pour les encodeurs MPEG-4 AVC pour lesquels la taille des blocs peut varier.

La solution proposée dans ce cas est d'appliquer le même principe mais sur une grille de 4 pixels par 4 pixels pouvant ainsi sélectionner toutes les frontières. En revanche, l'échelle finale de degré de visibilité devra être re-estimée avec les tests subjectifs. En effet, si le nombre de frontières à étudier sur l'image augmente, cela diminuera la moyenne obtenue.

L'autre possibilité est de considérer que comme les petites tailles de bloc sont associées aux objets en mouvement, ces objets sont compensés en mouvement et l'erreur d'estimation rendra les artefacts de bloc moins perceptibles. Dans ce cas, on garde juste une estimation sur des frontières positionnées à une fréquence de 8 pixels par 8 pixels qui sont censées être les plus visibles.

Ces cas nécessitent une étude particulière pour un domaine d'application plus large de la métrique.

4.1.2 Le changement de résolution d'une image

Dans le domaine de la télévision, de part le nombre important de formats d'affichages, nous pouvons souvent être confrontés à un changement de résolution de l'image et les tailles de blocs ne sont plus dans ce cas de 8 pixels par 8 pixels. Dans ce cas, la solution la plus envisageable est de faire une détection des tailles de blocs sur toute une image à l'instant t et de supposer que l'image à l'instant $t+1$ a ces mêmes tailles de bloc. Le BLE proposé a été effectué pour s'adapter à toutes tailles de blocs, m et n représentant respectivement la hauteur et la largeur d'un couple de blocs.

Cet aspect de position de frontières est souvent reproché [Winkler et al., 2001] aux métriques mais c'est un aspect difficile à prendre en compte et pas forcément utile pour toutes les applications. De plus, dans la majorité des systèmes, ce redimensionnement ou ce décalage n'intervient que pour l'affichage et donc après ces traitements. L'étude est malgré tout intéressante car elle permet d'appréhender les différentes tailles de MBs pour le standard MPEG-4 AVC.

4.2 L'aspect temporel

Nous ne prenons pas du tout en compte dans notre métrique l'aspect temporel. Cet aspect est également souvent reproché dans les articles de comparaison de métriques [Winkler et al., 2001], [Leontaris et Reibman, 2005], et ce reproche est totalement fondé si les méthodes utilisent une modélisation la plus complète possible des caractéristiques du SVH.

Dans notre cas, comme nous souhaitons utiliser la métrique comme un paramètre de contrôle de correction, nous préférons utiliser la métrique image à image d'une séquence et étudier la courbe de variation de la métrique. Ce genre d'informations nous renseigne sur les changements de scènes et sur une estimation moyenne de la qualité de la séquence au regard de la courbe. Néanmoins, une étude de l'impact de la dimension temporelle serait nécessaire pour l'estimation de la qualité des images compressées car les frontières de blocs ne sont pas perceptibles de la même façon qu'elles soient en mouvement ou statiques.

Dans le cadre de comparaison de métriques, le BLE fonctionnera car cette étude en fonction du pire cas sera la même pour toutes les méthodes. Un re-calibrage avec des tests subjectifs sur des séquences pourraient changer l'échelle si nécessaire.

4.3 Le masquage de luminance et de la chrominance

Dans cette métrique, nous avons tenu compte du masquage d'entropie en évaluant les variations d'informations à l'intérieur des blocs. En revanche, nous n'avons pas tenu compte du masquage en luminance. Nous avons négligé cet aspect car ce masquage se devait aussi de prendre en compte la chrominance et nécessitait un changement d'espace. Sans ces deux aspects liés à la perception, la métrique reste bien corrélée à notre perception mais il est tout à fait possible de tenir compte de ces caractéristiques avant d'estimer la visibilité des blocs. Cela nécessite une étape de pré-traitement qui peut venir se greffer à la métrique.

Le BLE est une métrique de l'effet de bloc corrélée à notre perception pour des images fixes. Pour les vidéos, elle peut être utilisée pour faire une étude comparative d'efficacité d'algorithmes mais les degrés de sensibilité ne seront pas totalement liés à l'échelle de perception définie à l'aide des résultats des test subjectifs. Néanmoins, les résultats donnent une bonne approximation. Pour les phénomènes de masquage de luminance et de couleur, la métrique peut être effectuée après un pré-traitement si cela est nécessaire et voulu. De même, pour des tailles de bloc différentes de 8 pixels par 8 pixels, la métrique s'adapte automatiquement si l'information de taille de bloc lui est donnée en paramètre d'entrée.

Dans ce chapitre, nous avons étudié les points forts et les points faibles des métriques de l'effet de bloc sans référence.

A partir de cette étude, nous avons repris la structure de ces métriques et nous avons introduit de manière algorithmique les phénomènes observés sur notre perception de l'effet de bloc pour aboutir au développement du BLE.

La corrélation avec des tests subjectifs et la comparaison de ces résultats avec d'autres métriques nous ont permis de valider le BLE pour des images de contenus différents et plus ou moins compressées.

Nous avons ensuite évoqué les limitations du BLE qui ne sont pas directement liées à notre application mais qui sont à prendre en compte pour le rendre plus autonome ou adaptable à la norme MPEG-4 AVC.

Nous avons donc mis en évidence la capacité du BLE à différencier la qualité des corrections en terme de suppression de l'effet de bloc. Sur la figure 4.27, le BLE montre que sur l'image de droite, l'effet de bloc n'est plus perceptible. Malgré ces résultats donnés par le BLE, on remarque sur la figure 4.27 que la qualité globale de perception de l'image corrigée est plus mauvaise que celle de l'image compressée : le BLE n'est donc pas suffisant. Nous proposons de compléter cette analyse dans le chapitre suivant avec l'estimation de la perception du flou.

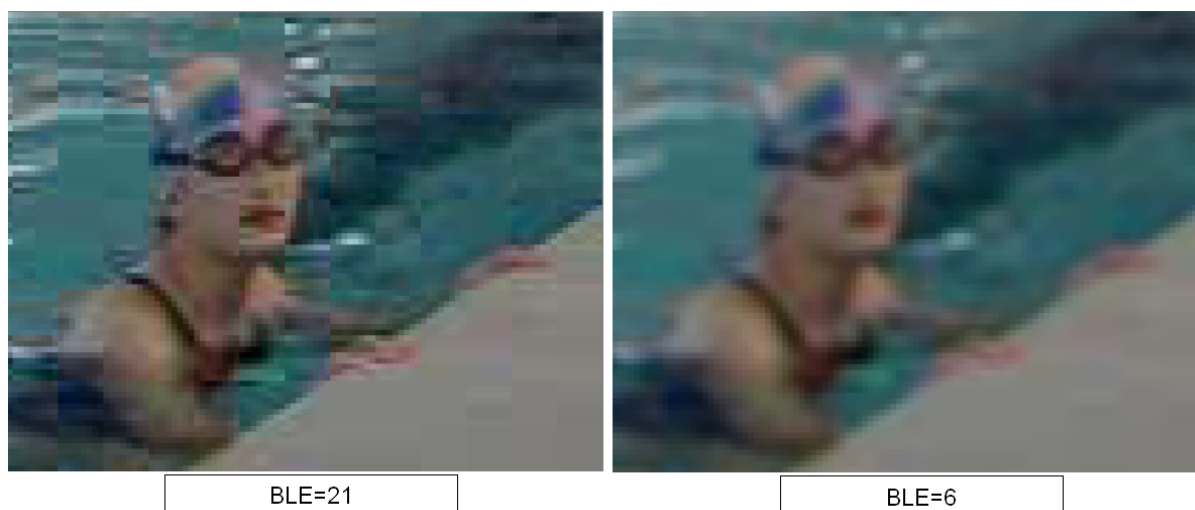


FIG. 4.27 – Le BLE indique que l'effet de bloc est gênant sur l'image compressée de gauche et que l'effet de bloc est imperceptible sur l'image floue de droite.

Chapitre 5

La mesure du flou

Nous cherchons à estimer la gêne occasionnée par le flou introduit par une correction. Néanmoins, le flou peut aussi être un artefact de compression ou encore la conséquence d'une défocalisation dans les domaines de la photographie. Nous commencerons donc ce chapitre par un état de l'art des métriques de flou dans ces différents domaines d'application. Nous ciblerons ensuite les métriques de flou sans référence pour pouvoir les exploiter dans notre contexte. Puis, à partir de la description de phénomènes liés à la discrimination du flou par le SVH, nous développerons une métrique de flou sans référence appelée BluM (Blur Metric). L'efficacité de ce BluM sera ensuite validée par des tests subjectifs et par comparaison aux résultats d'autres métriques. Enfin, bien que notre application nous limite à réaliser cette métrique uniquement sur la composante de luminance, nous finaliserons notre analyse par l'étude du rôle de la couleur dans la perception du flou.

1 Etat de l'art des différentes mesures de flou et de leurs applications

En terme de qualité d'images, le critère de netteté est l'un des plus recherché. Lorsqu'une image ne nous paraît pas nette, elle est floue, et ce flou peut être causé par :

- une manipulation matérielle : par exemple une mauvaise focalisation de la visée d'un photographe ou un mouvement de caméra.
- un traitement : par exemple un taux de compression trop élevé, une correction trop agressive ou une mauvaise interpolation censée créer de l'information pour agrandir une image.
- une création artistique, c'est à dire volontairement créé.

Pour les deux premières applications, le flou est un artefact et la littérature propose des analyses du flou dans ces différents domaines. Nous commencerons par décrire ces méthodes. Le flou artistique n'étant pas un artefact, nous n'en tiendrons pas compte.

1.1 Le flou causé par une mauvaise focalisation

Dans le domaine de la photographie, la focalisation consiste à obtenir une image nette d'un objet. Lors de l'acquisition d'une image ou d'une vidéo, si un objet se trouve hors du plan focal du système optique de l'appareil photo ou de la camera, il paraîtra flou sur le support restitué. Nous allons décrire deux catégories méthodes utilisées pour détecter ce flou.

1.1.1 La modélisation du flou

La première catégorie de méthode utilise une modélisation du flou pour estimer sa gêne.

Principe Dans des conditions de modélisation géométrique avec une lentille simple, l'image floue d'un point source est associée à une tâche circulaire d'intensité constante dont la fonction d'étalement ou PSF (Point Spread Function) déterminera le degré de flou. Dans la pratique, pour modéliser les pertes, la PSF est assimilée à une fonction gaussienne. Ainsi, une image floue sera modélisée par une image convoluée avec une gaussienne [Pentland et al., 1989].

Dans [Elder et Zucker, 1998] et [Hu et Haan, 2006], les auteurs utilisent la fonction gaussienne donnée par l'équation suivante :

$$g(n, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad n \in Z \quad (5.1)$$

où σ est l'inconnu à déterminer donnant la force du flou causé par la dé focalisation.

Résultat La figure 5.1 tirée de [Hu et Haan, 2006] représente la détection du flou par les méthodes de [Elder et Zucker, 1998] et [Hu et Haan, 2006] sur l'image *Léna*. Plus les pixels sont clairs, plus la zone est floue. La seconde méthode (troisième image) a tendance à mieux distinguer le flou dans l'arrière plan. Néanmoins, ce flou est un flou voulu par le photographe qui souhaite que l'on focalise notre attention uniquement sur *Léna* et pas sur l'arrière plan. De plus, ces méthodes ont tendance à définir les zones avec du détail comme les zones les plus nettes. Ce n'est pas toujours le cas car une zone avec très peu de détails peut également paraître nette.



FIG. 5.1 – La détection du flou sur *Léna* par les méthodes de [Elder et Zucker, 1998] et [Hu et Haan, 2006] : les pixels clairs indiquent une zone floue

1.1.2 L'utilisation des coefficients DCT

Au lieu de modéliser le flou, des méthodes proposent d'utiliser les coefficients de la DCT.

Principe Dans le chapitre 2 sur la description des artefacts, nous avons montré que le flou était un artefact de compression causé par la quantification des coefficients hautes fréquences. Marichal et al [Marichal et al., 1999] proposent d'analyser les coefficients de la DCT pour détecter le flou dû à un mouvement de la caméra ou à un objet hors de la focale. Leur méthode consiste à analyser l'occurrence des coefficients DCT non nuls de chacun des blocs de l'image.

Résultat En réalisant cette étude, les zones nettes avec peu d'informations seront encodées avec beaucoup de coefficients AC nuls et pourront être prises pour du flou. De plus, l'autre inconvénient de cette méthode est le risque de confusion entre le flou causé par la compression et le flou causé par une manœuvre matérielle.

1.2 Le flou gênant pour le SVH

Pour éviter ce genre de confusion et parce que notre intérêt est d'estimer la qualité des images diffusées, nous nous sommes concentrés sur les méthodes qui n'accordent pas d'importance à la cause et qui cherchent juste à détecter le flou gênant pour le SVH.

1.2.1 La conséquence du flou sur les contours

Ces méthodes cherchent à modéliser les conséquences du flou sur une image et plus particulièrement sur les contours des images : plus les contours sont diffus, moins l'image est nette. Ces méthodes, [Caviedes et Gurbuz, 2002], [Marziliano et al., 2002], [Marziliano et al., 2004], [Ong et al., 2003] sont divisées en une étape de détection de contours et une étape d'étude de ce contour.

La méthode proposée par Caviedes et al dans [Caviedes et Gurbuz, 2002] consiste à détecter et estimer les contours en calculant des statistiques avec le coefficient de Kurtosis sur les informations spatiales (les valeurs des pixels) et fréquentielles (les coefficients de la DCT). Ne souhaitant pas ré-estimer ces coefficients, nous n'exploiterons pas cette méthode. Dans la méthode de Ong et al [Ong et al., 2003][Ong et al., 2003], développée aussi sur la diffusion des contours, une importante partie de la validité de leur métrique est due au réglage des coefficients de leur formule

avec des tests subjectifs. Ces métriques étant liées aux images utilisées par les tests, nous n'utiliserons pas non plus cette méthode. En revanche, la méthode proposée par Marziliano et al dans *Marziliano et al* fait une estimation dans le domaine spatial et semble être la plus adaptée à notre contexte. Nous allons donc détailler cette méthode pour analyser comment les auteurs tiennent compte de la diffusion des contours.

1.2.2 La description de la méthode *Marziliano et al*

Principe Le diagramme bloc de cette méthode est représenté sur la figure 5.2. Dans un premier temps, une détection de contours en utilisant un filtre de Sobel est établie. Puis, l'idée est de détecter les extremum de ces contours. Ainsi, la diffusion du contour sera représentée par la distance entre ses extremum. La figure 5.3 illustre ce principe : les contours détectés (P1 et P3) sont indiqués par les pointillés en gras et les extremum correspondants (P2, P2' et P4, P4') sont indiqués par les pointillés les plus fins. La distance entre les extremum correspond au nombre de pixels qui les sépare.

Enfin, pour obtenir la valeur finale, toutes les distances sont sommées et normalisées par rapport au nombre total de contour.

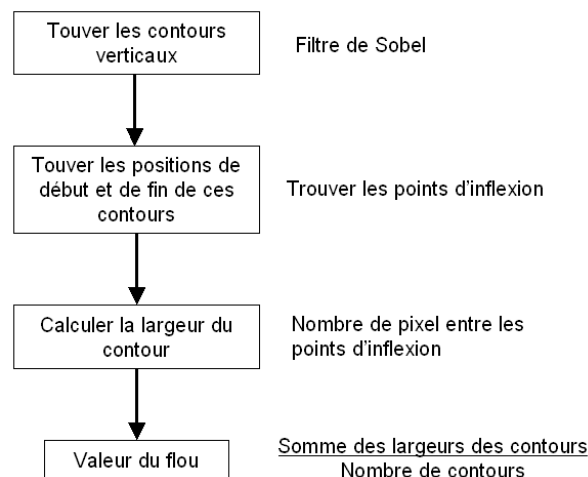


FIG. 5.2 – Diagramme bloc de la méthode de Marziliano et al [Marziliano et al., 2002].

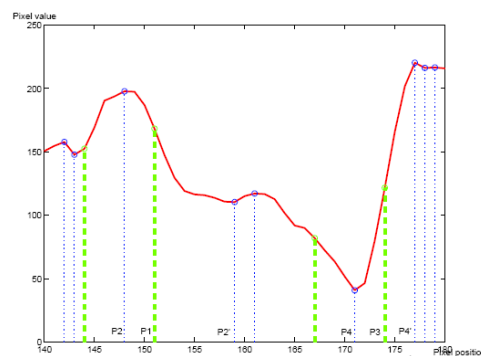


FIG. 5.3 – Principe de détection des points d'inflexion. Image tirée de [Marziliano et al., 2004]

Résultats Cette métrique donne des valeurs entre 0 (pas de flou visible) et 10 (flou très visible). L'efficacité de cette métrique est prouvée dans [Marziliano et al., 2002] et dans [Marziliano et al., 2004] par une corrélation des résultats avec d'autres métriques et avec des tests subjectifs sur des images rendues floues soit avec un filtre gaussien soit par la compression JPEG2000.

L'inconvénient Malgré son implémentation dans le domaine spatial, cette métrique contient une étape de recherche des extremum dont le coût d'implémentation n'est pas négligeable.

Pour évaluer le flou gênant pour le SVH, la métrique de Marziliano et al semble être la plus adaptée à notre contexte. Mais, en analysant ces métriques, nous avons observé un phénomène lié à notre discrimination du flou que nous avons traduit dans le développement d'une nouvelle métrique de flou, le BluM.

2 Proposition d'une métrique de flou : le BluM ou BlurMetric

Nous proposons d'expliquer les différentes étapes de cette méthode qui doit être sans référence et corrélée à notre perception. Pour ce faire, nous commençons par examiner notre capacité à discriminer différents niveaux de flou.

2.1 La discrimination du flou

En examinant le comportement du phénomène de flou sur des images plus ou moins floues, nous avons observé une limite au delà de laquelle nous n'étions plus capables de percevoir différents niveaux de flou.

Expérience de discrimination L'expérience représentée sur la figure 5.4 montre ce phénomène : lorsque l'on demande à plusieurs personnes de classer les images de la plus nette à la plus floue, les observateurs déterminent facilement que l'image la plus nette est l'image *b* et après des hésitations, quelques observateurs trouvent la suite du classement, c'est à dire l'image *c*, puis l'image *a*. Or, pour obtenir les images *b* et *c*, nous avons appliqué les traitements suivants :

- L'image *c* est l'image *b* floutée par un filtre moyenneur .
- L'image *b* est l'image *c* floutée par ce même filtre moyenneur.

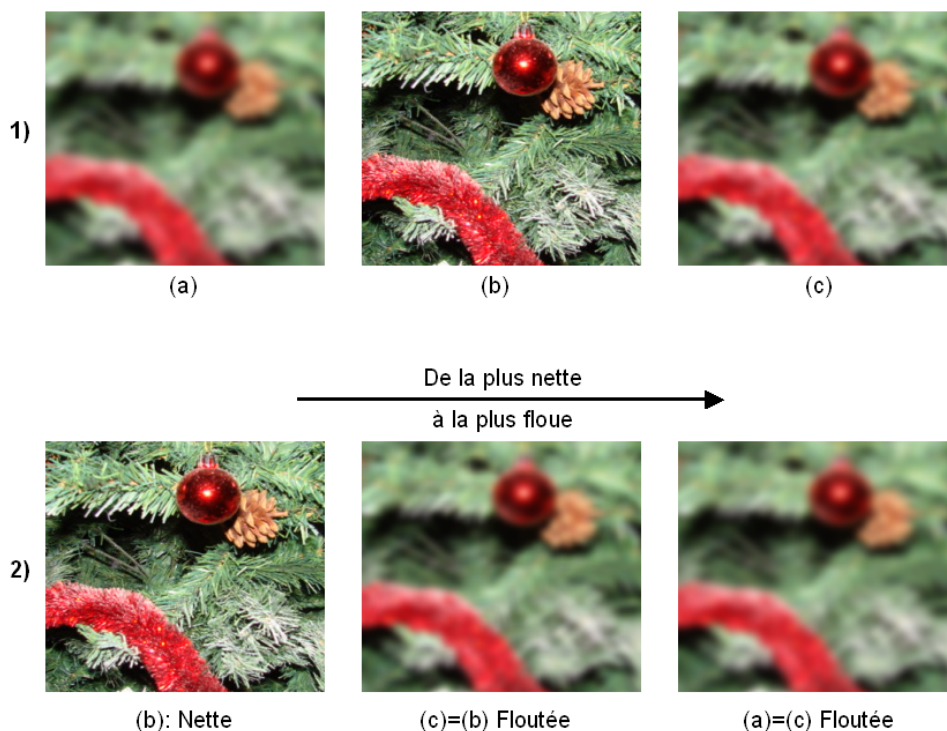


FIG. 5.4 – Expérience sur la discrimination du flou : 1) images à classer de la plus nette à la plus floue. 2) Résultat.

Conclusion de l'expérience De cette expérience, nous déduisons qu'à partir d'un certain niveau de flou, le SVH a des difficultés à distinguer deux niveaux de flou.

Explications du phénomène Ce phénomène s'explique en étudiant le comportement des pixels des images. En effet, plus on rend floue une image en appliquant de manière itérative un filtre passe bas, plus les valeurs des pixels d'une même région convergent vers une même valeur moyenne. Si on rend floue une image nette, les intensités des pixels voisins d'une même région vont varier significativement entre l'image nette et l'image floue. En revanche, si on rend floue une image déjà floue, comme les intensités des pixels voisins sont déjà proches, la variation d'intensité des pixels voisins de l'image rendue floue va être faible.

Ce phénomène est représenté sur le graphique de la figure 5.5. Cette figure représente les histogrammes des variations en valeur absolue des pixels horizontalement voisins. On remarque que le nombre de variations significatives diminue significativement de l'image nette à l'image floue et très peu de l'image floutée à l'image refloutée. Ces fortes variations d'intensité de l'image nette se trouvent transformées en faibles variations d'intensité comme le montre la partie gauche du graphique.

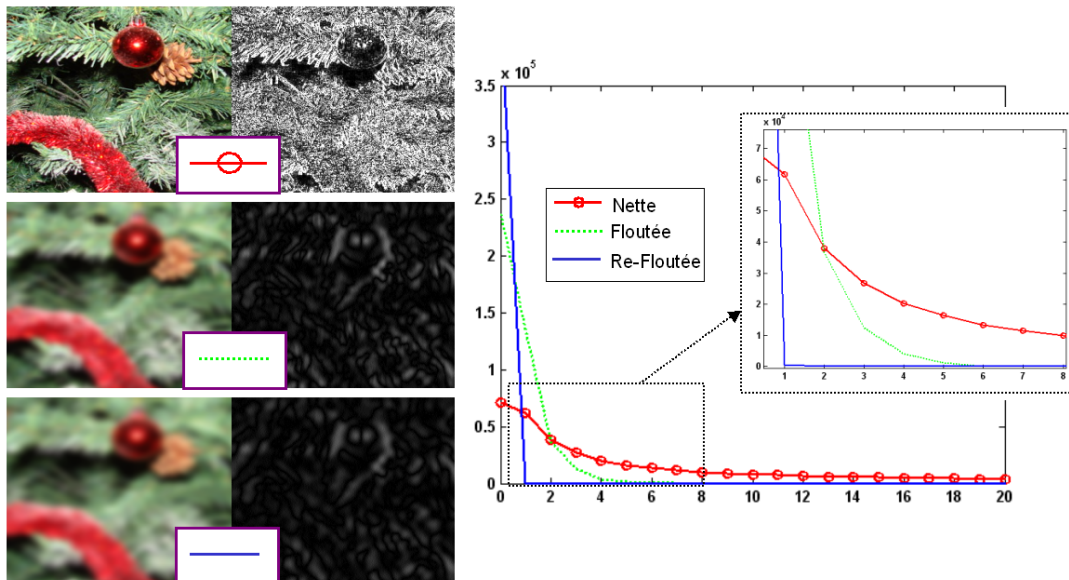


FIG. 5.5 – Histogrammes des variations entre pixels voisins sur une image nette, floutée et re-floutée.

2.2 Le principe de la métrique

L'idée clef de notre métrique est d'utiliser la propriété liée à notre discrimination du flou pour créer une image de référence. Le principe de cette métrique est représenté sur la figure 5.6 :

- Etape 1 : A partir d'une image dont le niveau de flou est à évaluer, la première étape consiste à générer une image très floue en appliquant un filtre assez agressif : cela permet d'obtenir une image floue quelque soit son contenu.
- Etape 2 : La deuxième étape s'applique sur l'image à estimer et l'image floutée. Sur ces deux images, nous étudions la distance en terme de différence absolue d'intensité entre pixels voisins.
- Etape 3 : La troisième étape consiste à analyser les variations d'une image à l'autre entre ces distances.

- Etape 4 : c'est l'étape finale qui permet de conclure sur le niveau de flou de l'image originale : si la variation est importante, cela signifie que l'image originale était nette. En revanche, si la variation est faible, cela signifie que l'image à la base était déjà floue.

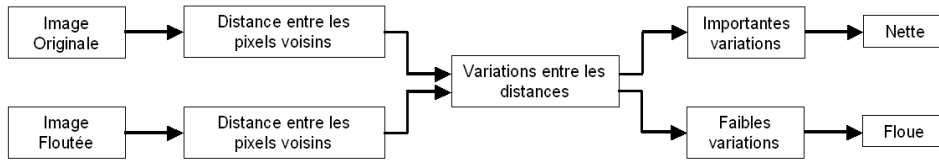


FIG. 5.6 – Principe de la métrique.

Ce principe décrit un état binaire (net ou flou) de l'image et un exemple de chaque cas est illustré sur les figures 5.7 et 5.8.

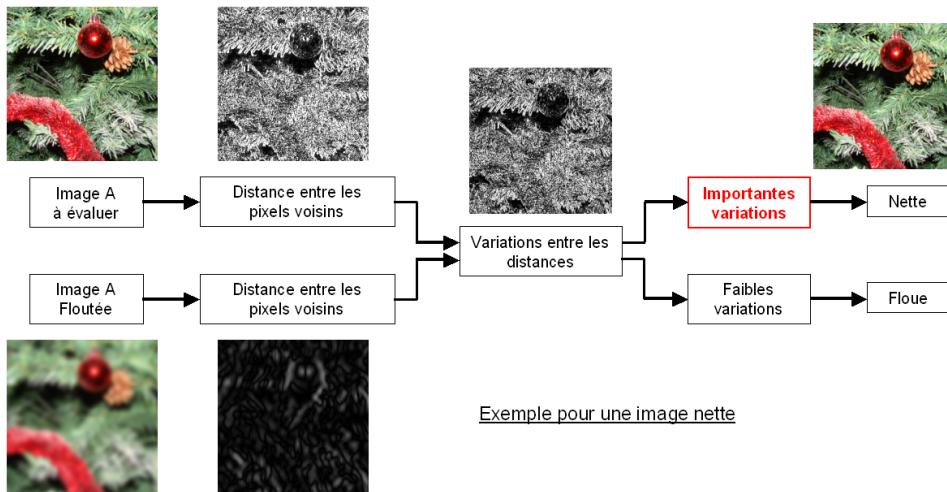


FIG. 5.7 – Illustration du principe de la métrique sur une image nette.

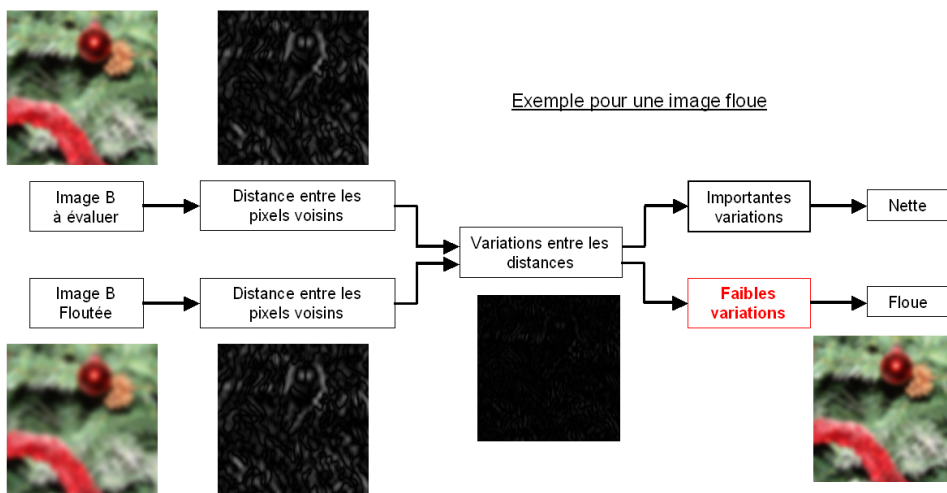


FIG. 5.8 – Illustration du principe de la métrique sur une image floue.

Avant d'expliquer comment nous avons intégré ce principe dans un algorithme, nous proposons de lister les autres caractéristiques liées à notre perception du flou qui nous ont guidées dans nos choix de développement.

2.3 Les caractéristiques à prendre en compte

Le contenu de l'image Le contenu de l'image joue un rôle important dans la perception du flou notamment pour les zones homogènes. En effet, l'image présentée sur la figure 5.9 nous paraît floue alors qu'en réalité, seulement 10% de l'image est floue. Même si la notion de flou est liée aux hautes fréquences et donc à la présence de variations d'information, les zones homogènes d'une image ne contenant pas de variations d'informations nous paraissent nettes. Ce phénomène est à prendre en compte dans la métrique.



FIG. 5.9 – La perception du flou suivant le contenu de l'image.

La perte et la génération d'information Sur la figure 5.10, nous avons représenté une lettre noire sur fond blanc rendue floue avec un filtre passe-bas. Le zoom sur les détails de ces images montrent que :

- Les intensités des pixels présents sur l'image nette ont diminué sur l'image floue. Cela se traduit par une perte de variation d'information.
- Des pixels d'intensité non nulle sont apparus sur l'image floue. Cela se traduit par une création de variation d'information au niveau de l'image floue qui n'existait pas sur l'image nette.

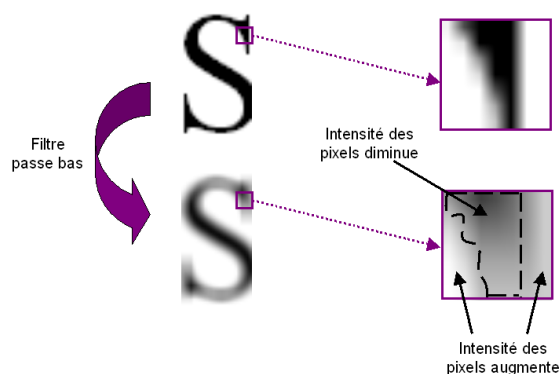


FIG. 5.10 – Les pertes et les gains des intensités des pixels après avoir flouté d'une image.

Ces deux aspects sont à prendre en compte si on veut comparer les variations entre des distances d'une image à l'autre. En effet, pour tenir compte du flou, nous cherchons à quantifier la perte des hautes fréquences. C'est pourquoi nous ne tiendrons pas compte pour la métrique des variations d'information créées par l'étape de filtrage des hautes fréquences.

Le flou directionnel Dans notre principe, nous souhaitons étudier la variation d'intensité entre pixels voisins. L'image représentée sur la figure 5.11 a été floutée uniquement horizontalement donc seules les variations entre pixels horizontalement voisins sont faibles. Pourtant, lorsque nous avons à la juger, nous la qualifions d'image floue et non d'image nette dans toutes directions sauf l'horizontale. Nous aurons à tenir compte de cette remarque dans notre métrique.



FIG. 5.11 – La perception du flou unidirectionnel.

2.4 L'implémentation de la métrique

Avant de décrire l'algorithme, nous allons définir son domaine de définition.

2.4.1 Le domaine de définition

Pour des raisons de coût d'implémentation et parce que nos composantes de chrominances sous-échantillonnées sont parfois inexploitable, nous estimons le flou uniquement sur la composante de luminance de l'espace YCbCr. Nous justifierons dans la suite que l'étude du flou sur la luminance est suffisante.

Nous privilégions deux directions de recherche pour le flou, la direction horizontale et la direction verticale car, lorsqu'il s'agit d'un flou de bougé, ce flou est fréquemment généré dans l'une de ses deux directions (mouvement de caméra). Néanmoins, l'algorithme présenté ci après peut s'adapter à toutes les directions.

2.4.2 La description de l'algorithme

Le diagramme bloc de la figure 5.12 présente les étapes que nous allons décrire précisément.

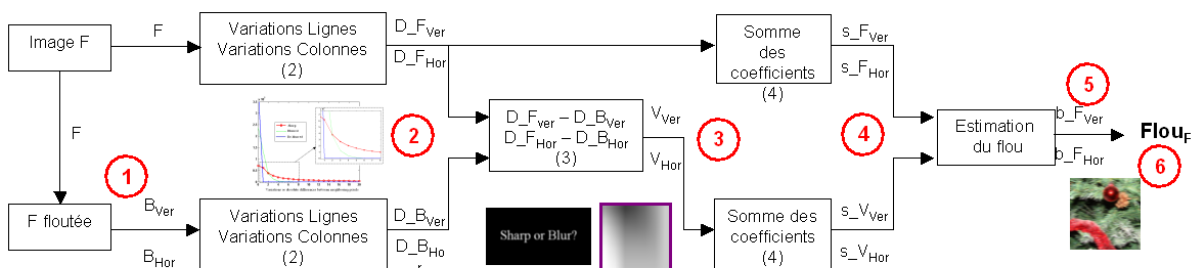


FIG. 5.12 – Description des étapes de l'algorithme de la mesure du flou.

Etape 1 Soit F la composante de luminance d'une image de taille $m \times n$ pixels où m représente le nombre de lignes et n le nombre de colonnes. Cette première étape consiste à rendre floue cette image de manière à obtenir une image floue quelque soit son contenu. De ce fait, le filtrage doit être assez fort. Pour créer les deux images B_{Ver} et B_{Hor} floutées respectivement dans les directions verticales et horizontales, nous choisissons des filtres moyennneur à une dimension dont les noyaux sont constitués de 9 pixels et donnés par les formules suivante :

$$h_v = \frac{1}{9} \times [1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1] \quad h_h = transpose(h_v) = h'_v \quad (5.2)$$

$$B_{ver} = h_v * F \quad B_{hor} = h_h * F \quad (5.3)$$

Remarque 1 : nous avons testé plusieurs type de flou et à force égale, les résultats sont équivalents. Ce choix de filtre peut être modifié, l'important étant juste d'obtenir une image floue.

Remarque 2 : Nous supposons pour cette étude que nous estimons la qualité des images à une distance identique pour toutes les images car la notion de distance va influencer notre perception du niveau de flou. En revanche, il est possible d'imaginer une variable mettant en relation la taille du noyau du filtre utilisé en fonction de la distance d'observation : plus la distance augmente, moins l'image nous paraîtra flou, donc plus on augmentera la taille du noyau pour que la perception du flou reste la même.

Etape 2 Cette étape définit la notion de distances entre deux pixels voisins. Pour étudier les distances entre pixels voisins, nous choisissons d'appliquer la norme L_1 dans les deux directions. Ce choix a été influencé par la contrainte d'implémentation à faible coût. Nous obtenons ainsi les images distances D_{FVer} et D_{FHor} de l'image (composée de m lignes et de n colonnes) à estimer et les images distances D_{BVer} et D_{BHor} de l'image rendue floue :

$$D_{FVer}(i, j) = Abs(F(i, j) - F(i - 1, j)) \quad \text{pour } i = 1 \ m - 1, j = 0 \ n - 1 \quad (5.4)$$

$$D_{FHor}(i, j) = Abs(F(i, j) - F(i, j - 1)) \quad \text{pour } i = 0 \ m - 1, j = 1 \ n - 1 \quad (5.5)$$

$$D_{BVer}(i, j) = Abs(B(i, j) - B(i - 1, j)) \quad \text{pour } i = 1 \ m - 1, j = 0 \ n - 1 \quad (5.6)$$

$$D_{BHor}(i, j) = Abs(B(i, j) - B(i, j - 1)) \quad \text{pour } i = 0 \ m - 1, j = 1 \ n - 1 \quad (5.7)$$

Etape 3 Pour tenir compte du cas expliqué sur la figure 5.10, nous choisissons de n'étudier que les distances qui ont diminuées entre l'image à estimer et l'image floutée. Nous obtenons ainsi l'image variation des distances verticales V_{Ver} et l'image variation des distances horizontales V_{Hor} entre l'image à estimer F et l'image rendue floue B .

$$V_{Ver} = Max(0, D_{FVer}(i, j) - D_{BVer}(i, j)) \quad \text{pour } i = 1 \ m - 1, j = 1 \ n - 1 \quad (5.8)$$

$$V_{Hor} = Max(0, D_{FHor}(i, j) - D_{BHor}(i, j)) \quad \text{pour } i = 1 \ m - 1, j = 1 \ n - 1 \quad (5.9)$$

Etape 4 Puis, pour comparer ces distances obtenues sur l'image à estimer et sur cette même image que nous avons flouté, nous calculons la somme des distances initiales $s_{_FVer}$ et $s_{_FHor}$ dans l'image à estimer et la somme des variations calculées $s_{_VVer}$ et $s_{_VHor}$ dans l'image floutée.

$$s_{_FVer} = \sum_{i,j=1}^{m-1,n-1} D_{_FVer}(i,j) \quad \text{et} \quad s_{_FHor} = \sum_{i,j=1}^{m-1,n-1} D_{_FHo}(i,j) \quad (5.10)$$

$$s_{_VVer} = \sum_{i,j=1}^{m-1,n-1} V_{Ver}(i,j) \quad \text{et} \quad s_{_VHor} = \sum_{i,j=1}^{m-1,n-1} V_{Hor}(i,j) \quad (5.11)$$

Rmarque : Ces étapes de sommation peuvent être effectuées en même temps que les étapes 2 et 3, ce qui minimise le coût d'implémentation.

Etape 5 L'étape finale consiste à comparer ces sommes en les normalisant de la manière suivante :

$$b_{_FVer} = \frac{s_{_FVer} - s_{_VVer}}{s_{_FVer}} \quad (5.12)$$

$$b_{_FHor} = \frac{s_{_FHor} - s_{_VHor}}{s_{_FHor}} \quad (5.13)$$

Notons que les deux résultats obtenus sont toujours positifs ou nuls car ils contiennent que les variations qui ont diminuées. Ces valeurs sont comprises entre 0 et 1, 0 représentant la netteté et 1 le flou.

Etape 6 Enfin, pour tenir compte de la remarque faite sur la perception du flou de mouvement 5.11, nous choisissons le flou le plus visible comme flou final.

$$blur_{_F} = Max(b_{_FVer}, b_{_FHor}) \quad (5.14)$$

L'échelle du BluM Nous venons de proposer une mesure sans référence du niveau de flou sur une image. Cette métrique est normalisée entre 0 et 1 suivant l'échelle de perception représentée sur la figure 5.13.



FIG. 5.13 – Echelle de perception du BluM.

Nous venons de développer une métrique sans référence d'estimation de la gêne occasionnée par le flou sur une image. Ce BluM tient compte de plusieurs phénomènes liés à notre perception du flou. Nous allons vérifier que ces résultats sont valides et corrélés avec notre perception.

3 Validation du BluM et résultats

La validation de la métrique s'effectue en trois temps. Tout d'abord, une série de tests subjectifs nous permet d'analyser le comportement des observateurs pour noter le niveau de flou. Puis, nous confronterons ces notes attribuées aux notes données par notre métrique. Enfin, nous comparerons la fiabilité de cette métrique avec celle de *Marziliano et al* [Marziliano et al., 2002].

3.1 Les tests subjectifs

Ces tests ont été effectués en collaboration avec Thierry Dolmière [Dolmière, 2006] et en suivant les recommandations de l'ITU [BT.500-11, 2002]. Nous commencerons par décrire le protocole utilisé.

La procédure Une première session de 10 images permet de calibrer le SVH sur les différents niveaux de dégradation. Puis, le test contient deux parties équivalentes séparées par un entracte de 3 minutes minimum. Chaque partie contient 132 images. Chaque image est présentée pendant une durée de 5 secondes mais l'utilisateur peut réduire ce temps s'il est sûr de son jugement. L'échelle de notation lui est alors présentée et il a le temps qu'il souhaite pour noter. Après la notation, une image grise est affichée pendant 0.5 secondes pour éviter le phénomène de rémanence rétinienne. L'affichage des images est réalisé dans un mode aléatoire.

La salle Afin d'éviter toute influence de l'environnement sur l'opinion du sujet, la salle choisie représentée sur la figure 5.14 a été montée dans un environnement gris neutre à l'aide de rideaux. L'ITU recommande d'éclairer la salle avec un illuminant D65 (6500K) pour fatiguer les yeux le



FIG. 5.14 – La salle des tests subjectifs.

moins possible. N'ayant pas de tels moyens, nous avons choisi de faire le noir total pendant les expériences. Tout voyant lumineux étaient cachés. Les rideaux pendus évitaient que l'écran ne se réfléchisse sur le plafond, les murs et la table.

L'écran L'écran utilisé pour les expériences est un écran CRT *21 View Sonic G220f* [Sonic,] : les réflexions sur cet écran sont minimisées. Pour nos expériences, nous l'avons calibré à 6500K ($\pm 100K$) à l'aide du GretagMacBeth Eye-One. La résolution de l'écran a été fixée en 1400x1050 dans le but de correspondre avec la taille des images présentées. Pour chaque expérience, la distance d'observation de l'écran a été adaptée suivant sa diagonale comme cela est recommandé dans [BT.500-11, 2002]. L'écran a été positionné de façon à ce que la tête des sujets soit dans l'axe normal par rapport à sa surface.

La manette Nous avons utilisé une manette de jeu *APM* pour faire passer les tests. Celle-ci a permis d'accélérer le rythme des expériences en faisant diminuer la durée des tests de moitié. Nous avons remarqué que pour nos sujets, une procédure à suivre avec une manette est plus intuitive qu'avec un clavier. Elle permet de simplifier le protocole expérimental et d'éviter des erreurs.

Les images Le lot d'images utilisé était constitué de 12 images naturelles aléatoirement filtrées pour représenter différents niveaux de dégradation. Nous avons utilisé des filtres Gaussien de noyaux 3x3 et 5x5, des filtres moyenneur de noyaux 3x3, 5x5, 7x7 et 9x9, des filtres moyen directionnel de noyaux 3x1, 5x1, 7x1 et 9x1 et un filtre disque Gaussien de diamètre 11 pixels. Les 12 images d'origine représentent des scènes naturelles plus ou moins texturées. La figure 5.15 montre un échantillon de ces 132 images. Ces images, présentées dans un ordre aléatoire, ont été évaluées une fois par partie. L'expérience comportant deux parties permettait donc d'obtenir deux évaluations par sujet pour une même image.



FIG. 5.15 – Echantillon des images présentées aux observateurs.

Les observateurs La perception de la qualité d'une image peut varier, notamment dans la façon qu'à le sujet d'observer une image. Une personne expérimentée pourra repérer plus rapidement certains défauts qu'une autre. Pour ces raisons, nous avons distingué les personnes étant débutantes dans le domaine (aucune notion de traitement d'images), novices (exerçant un métier en rapport avec l'image) et expérimentées (notions avancées de traitement d'images). Ces dernières faisaient partie du personnel du laboratoire GIPSA.

Le jugement des observateurs Contrairement à la précédente étude subjective sur l'effet de bloc, nous avons demandé aux observateurs de juger la qualité générale des images et non pas de juger uniquement le flou. Les images choisies étant sans artefacts de compression mais plus ou moins floutées, les résultats de ces tests sont donc directement liés à la notion de flou.

3.2 L'analyse des tests subjectifs

Ces tests subjectifs représentent une étape longue à mettre en place mais qui donnent des renseignements intéressants sur le comportement des observateurs.

3.2.1 L'analyse du comportement des observateurs

En suivant les recommandations de l'ITU, le calcul du MOS et l'analyse statistique ont permis d'éliminer 3 des 15 observateurs, dont les réponses n'étaient pas cohérentes avec le reste du groupe. Puis, nous avons étudié les temps mis par les observateurs pour juger la qualité du flou sur une image. Ces temps sont représentés sur la figure 5.16.

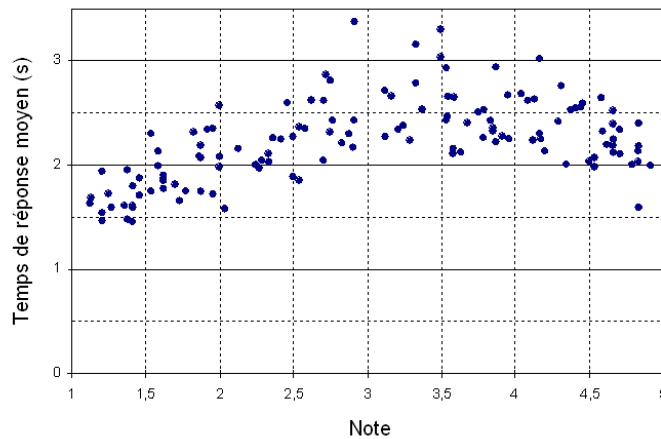


FIG. 5.16 – Temps de réponse moyen pour l'échelle des notes.

Nous observons 3 phénomènes importants :

- Il y a très peu d'hésitations pour les images très floues dont la note est faible. Les temps utilisés pour noter sont courts.
- En moyenne, les temps les plus longs sont ceux pris pour donner des notes moyennes. Lorsque l'image est de mauvaise qualité, la note est rapidement donnée, lorsqu'elle est de bonne qualité, la note est donnée avec une moyenne plus grande, le temps d'être sûr qu'il n'y pas de défaut. Enfin, lorsque l'image est de qualité moyenne, les observateurs cherchent les défauts et évaluent leur importance par rapport à ce qu'ils ont vu ce qui nécessite un temps plus grand.
- Il y a un léger saut psychologique dans l'évaluation du flou. En effet, il y a moins de scores entre 2 et 3.5, alors que les images étaient floutées de façon progressive. Ce phénomène signifie que la perception du flou est proche d'un état binaire : nette ou flou. Cela est corrélé avec les temps de note et avec l'aspect de la perception que nous avons utilisé pour établir la métrique.

3.2.2 La corrélation du test subjectif avec notre métrique de flou

Après avoir récolté ces résultats, nous les avons confrontés aux valeurs que donnait la métrique pour les mêmes images que celles du test subjectif. Ces résultats sont représentés par des triangles sur la figure 5.17. L'intervalle de confiance à 95% est également indiqué.

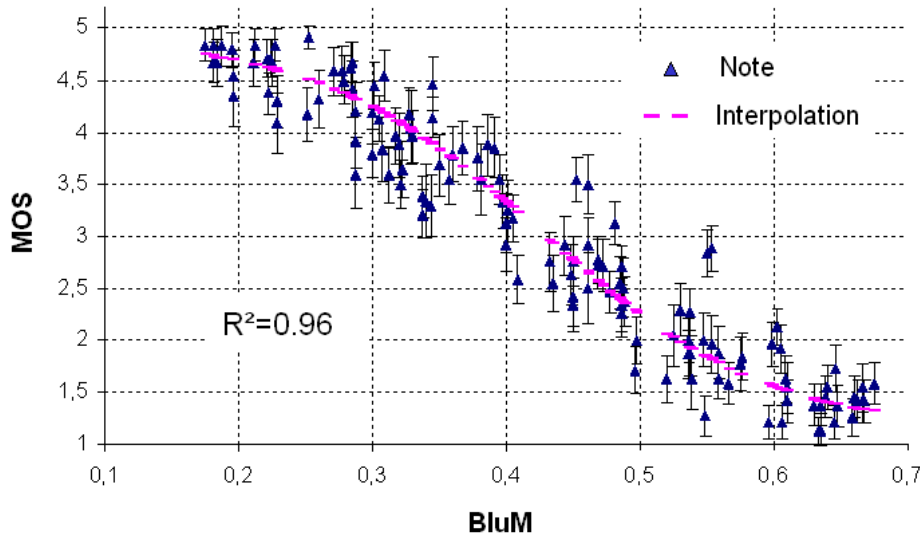


FIG. 5.17 – Corrélation entre le MOS et le BluM.

A première vue, nous observons une bonne corrélation entre la métrique et les tests subjectifs. Pour évaluer plus précisément cette corrélation, nous avons cherché une relation entre ces deux types de jugements. Nous avons suivie l'approche proposée par l'ITU qui consiste à faire une approximation par une fonction logistique symétrique : tenant compte du fait que les résultats découlent d'une expérience psychophysique et du saut psychologique observé, nous cherchons à faire correspondre les résultats objectifs et subjectifs à une courbe psychométrique symétrique de type sigmoïde. Si A_{Obj} représente les valeurs de la métrique et A_{sub} les valeurs obtenues par les tests subjectifs, la courbe sigmoïde est de type :

$$A_{Sub} = f(A_{Obj}) = \frac{A}{1 + \exp(G \times (A_{Obj} - D_M))} + B \quad (5.15)$$

où A et B sont des constantes et G et D_M sont les variables à déterminer pour obtenir la meilleure interpolation. Pour ce faire, nous utilisons les propriétés de f :

La fonction f est dérivable dans l'espace des réels car :

$$\forall A_{Obj}, 1 + \exp(G \times (A_{Obj} - D_M)) \neq 0 \quad (5.16)$$

La dérivée de f est alors :

$$f'(A_{Obj}) = \frac{-A \times G \times \exp(G \times (A_{Obj} - D_M))}{(1 + \exp(G \times (A_{Obj} - D_M)))^2} \quad (5.17)$$

La fonction f dépend du signe de G et est strictement monotone. On en déduit ses limites :

$$\begin{cases} G < 0 & G > 0 \\ \lim_{A_{Obj} \rightarrow -\infty} f(A_{Obj}) = \min(f(A_{Obj})) = B & \lim_{A_{Obj} \rightarrow -\infty} f(A_{Obj}) = \max(f(A_{Obj})) = A + B \\ \lim_{A_{Obj} \rightarrow +\infty} f(A_{Obj}) = \max(f(A_{Obj})) = A + B & \lim_{A_{Obj} \rightarrow +\infty} f(A_{Obj}) = \min(f(A_{Obj})) = B \end{cases} \quad (5.18)$$

Les extremum (min et max) correspondent aux notes extrêmes moyennes attribuées par les sujets lors des tests subjectifs. On recherche une interpolation décroissante, donc pour $G > 0$, on a :

$$\begin{cases} A + B = 4,92 \\ B = 1,13 \end{cases} \quad (5.19)$$

Puis, on calcule p , la variable relative aux notes subjectives :

$$p = \frac{f(A_{Obj}) - \min(f(A_{Obj}))}{\max(f(A_{Obj})) - \min(f(A_{Obj}))} = \frac{1}{1 + \exp(G(A_{Obj} - D_M))} \quad (5.20)$$

Ainsi, pour calculer les coefficients D et G_M , on pose :

$$I = \frac{1}{p} - 1 \quad \text{avec} \quad p = \frac{A_{Sub} - B}{A} \quad (5.21)$$

et on obtient

$$I = \exp(G \times (A_{Obj} - D_M)) \quad \text{et} \quad \ln(I) = G \times (A_{Obj} - D_M) \quad (5.22)$$

On connaît A_{Obj} et on peut calculer I pour chaque image. Ainsi, en traçant la courbe $\ln(I) = f(A_{Obj})$, on peut effectuer une approximation linéaire et trouver les coefficients G et D_M . La figure 5.18 illustre cette courbe et son approximation.

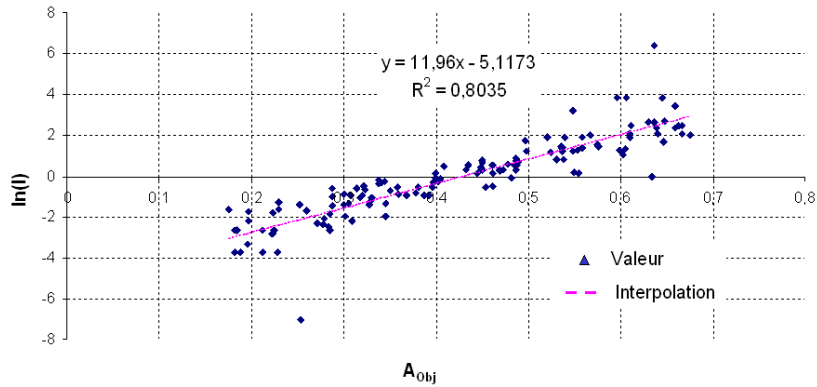


FIG. 5.18 – Approximation linéaire pour déterminer les coefficients G et D_M .

La droite obtenue par approximation nous permet d'obtenir les résultats suivants :

$$\begin{cases} \ln(I) = 11,96 \times A_{Obj} - 5,11 = G \times (A_{Obj} - D_M) \\ G = 11,96 \\ D_M = \frac{5,11}{11,96} \end{cases} \quad (5.23)$$

L'équation finale de la courbe d'interpolation est alors donnée par :

$$A_{Sub} = f(A_{Obj}) = \frac{3,79}{1 + \exp(11,96 \times A_{Obj} - 5,11)} + 1,13 \quad (5.24)$$

Cette interpolation est représentée sur la figure 5.17 en pointillé. Le coefficient de corrélation obtenu entre les valeurs subjectives et l'approximation de la courbe est calculée avec *Excel*. La valeur obtenue de 0.96, est proche de 1 ce qui signifie que la métrique est bien corrélée aux jugements des observateurs.

Cette équation nous permet de trouver une correspondance de degré de perception du flou à chaque valeur de la métrique. L'échelle simplifiée de ces données est représentée sur la figure 5.19. Pour qu'une image soit de qualité correcte en terme de flou, il faut que la métrique lui attribue une valeur environ < 0.35 .

Dégradation	Note subjective	Métrique de Flou
Imperceptible	5	0.2
Perceptible mais non gênant	4	0.35
Légèrement gênant	3	0.45
Gênant	2	0.55
Très gênant	1	0.7

FIG. 5.19 – Attribution d'un degré de visibilité du flou en fonction des valeurs de la métrique.

3.2.3 Les résultats

Sur différentes images Nous avons testé la métrique sur des images ne faisant pas forcément partie de la série utilisée pour les tests subjectifs. La figure 5.20 montre un échantillon de ces images associées à la valeur de la métrique. Ces images sont issues de sources différentes (photographie ou image de télévision) et de contenu différent (artificiel ou naturel) et on remarque que la métrique fait une bonne estimation du niveau de flou.

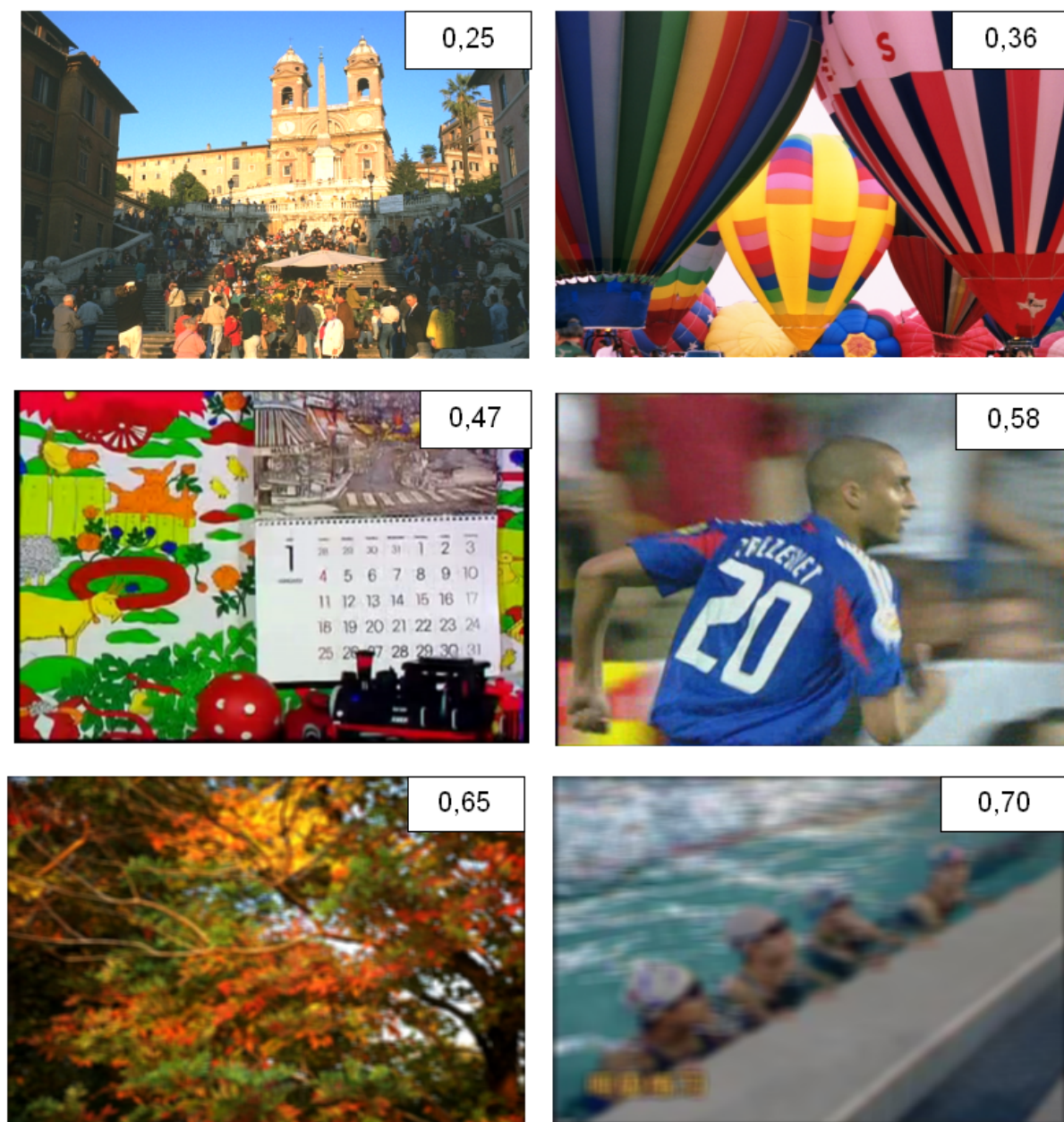


FIG. 5.20 – Résultats de la métrique de flou sur différentes images.

La gestion des différents contenus Nous avons fait remarquer dans l'étude du flou que les variations d'information pouvaient être influencées par le contenu de l'image. L'exemple illustré sur la figure 5.21 montre que pour deux images nettes de contenus totalement différents, la métrique fait une bonne estimation de la perception du flou.



FIG. 5.21 – Robustesse de la métrique selon le contenu.

Retour au problème initial Nous avons construit la métrique de flou sur le principe de notre limitation à percevoir le flou. Sur la figure 5.22, nous avons indiqué les valeurs données par la métrique de flou sur la série des 3 images prises en exemple pour montrer notre limitation. Nous pouvons remarquer que la métrique différencie à peine les deux images floues. En comparant à l'échelle de perception du flou, ces valeurs, bien supérieures à 0.7, montrent que ces images perçues sont très floues.

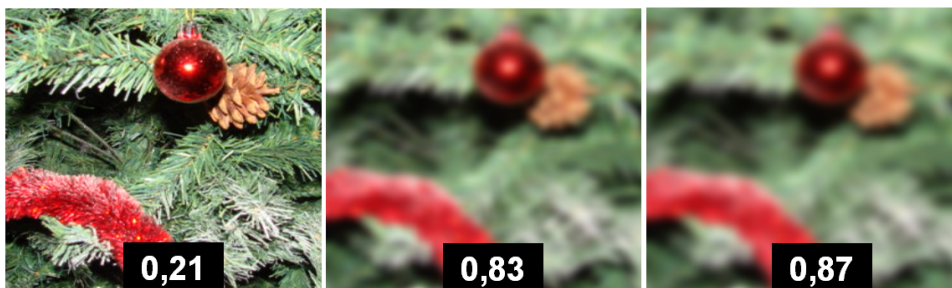


FIG. 5.22 – Corrélation de la métrique avec notre limite de perception du flou.

Ces exemples, associés aux résultats des tests subjectifs, confirment le fait que notre métrique de flou est corrélée à notre perception du flou.

3.3 La comparaison du BluM avec la métrique de *Marziliano et al*

A titre de comparaison, nous avons confronté les résultats des tests subjectifs à la métrique proposée dans [Marziliano et al., 2002]. La figure 5.23 illustre cette corrélation. Nous remarquons que leur corrélation est également bonne à quelques points près. Nous avons étudié cette corrélation de la même manière que pour notre métrique et nous avons trouvé l'interpolation suivante représentée par la courbe en pointillée :

$$A_{Sub} = f(A_{Obj}) = \frac{3,79}{1 + \exp(1,17 \times A_{Obj} - 6,89)} + 1,13 \quad (5.25)$$

Le coefficient de détermination déterminé avec Excel nous donne une valeur de 0,88 qui est légèrement moins bonne que notre valeur de 0,96. Pour une analyse plus profonde, nous avons alors représenté sur la figure 5.24, les résultats obtenus pour les deux métriques suivant la catégorie de flou utilisée. Nous remarquons que notre métrique est toujours incluse dans l'intervalle de confiance des résultats des tests subjectifs, ce qui montre son efficacité. L'efficacité de la métrique de *Marziliano et al* est montrée pour tous les types de flou sauf pour les flous de bougé ou leur métrique est légèrement au dessus de l'intervalle de confiance. Cela s'explique par le fait que les auteurs ne prennent pas en compte le maximum des 2 flous et font une moyenne. Ainsi, pour des flous de bougé, ils sousestiment le flou et donc surestiment la qualité.

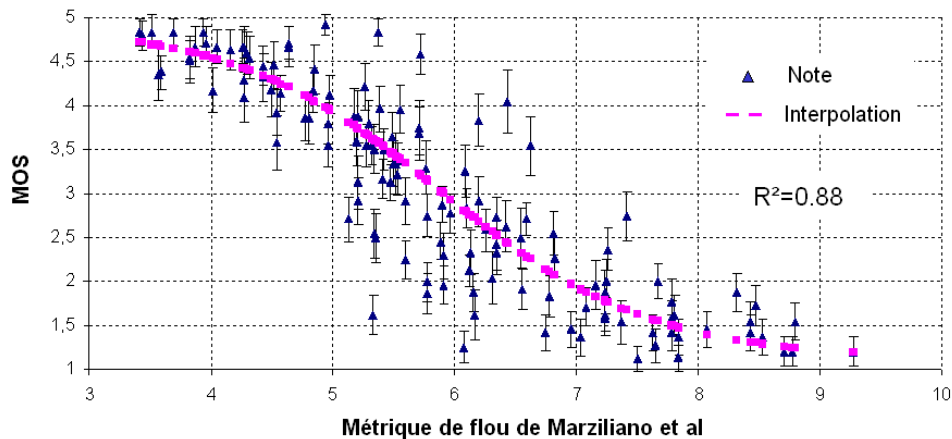


FIG. 5.23 – Corrélation de la métrique de *Marziliano et al* avec le MOS.

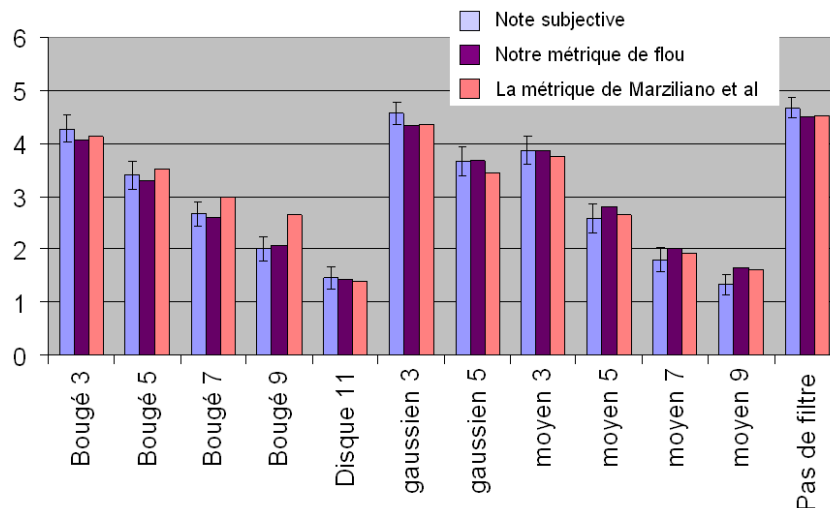


FIG. 5.24 – Corrélation entre les notes subjectives et le BluM et la métrique de *Marziliano et al* en fonction du flou utilisé.

3.4 La robustesse au bruit

Nous avons testé la robustesse de notre métrique au bruit analogique et au bruit numérique. Sur la figure 5.25, le premier couple d'image représente une image originale et cette même image que nous avons floutée. Nous observons que la métrique de flou fait une bonne estimation du niveau de flou.

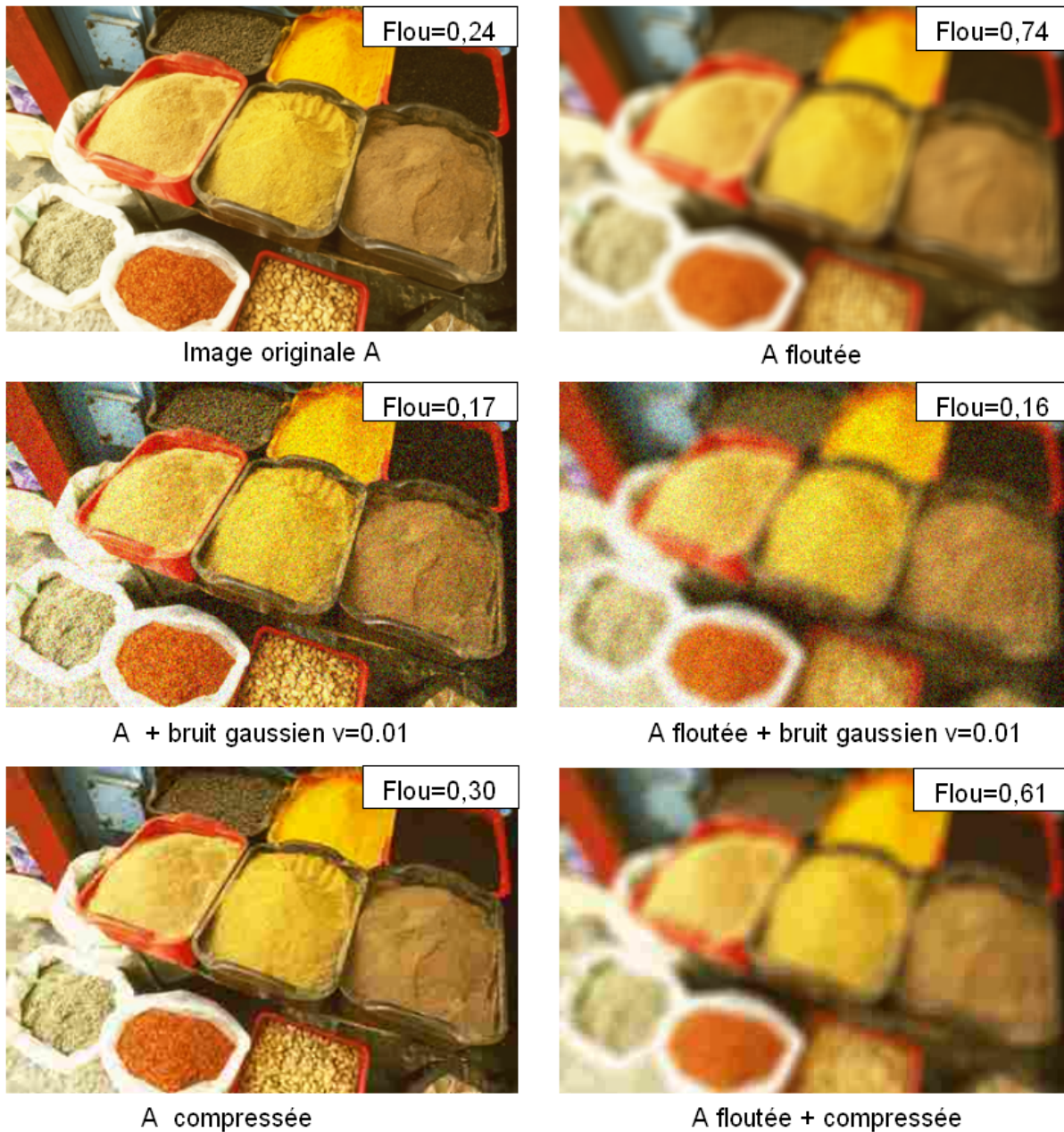


FIG. 5.25 – L'analyse de la robustesse de la métrique au bruit analogique et au bruit numérique.

Le bruit analogique Puis, nous avons soumis ce couple d'image à un bruit gaussien qui est une approximation du bruit analogique. Le résultat obtenu est une valeur très faible du niveau de flou pour les deux images, ce qui s'explique par le fait que le bruit ajoute de l'information qui renforce l'impression de netteté. En revanche, il n'est pas normal d'obtenir une note plus faible pour l'image floutée bruitée que pour l'image nette bruitée. Cela montre que la métrique ne fait pas une bonne estimation : en effet, en analysant les variations pixel à pixel, la métrique ne distingue que le bruit et n'est pas capable de distinguer l'arrière plan qui est l'information principale remarquée par le SVH. Face à du bruit analogique, notre métrique n'est donc pas corrélée à notre perception.

Pour comparer, nous avons également appliqué la métrique de *Marziliano et al* et obtenu 2.04 pour l'image nette bruitée et 1.88 pour l'image floue bruitée : la mesure de la diffusion du contour a été également faussée par le bruit gaussien rendant ainsi la métrique non corrélée à notre perception.

Les améliorations à envisager Deux piste sont possibles pour améliorer le BluM en le rendant robuste au bruit analogique :

- Le BluM n'est pas robuste au bruit analogique à cause de son étude pixel à pixel. On pourrait donc envisager un calcul des distances entre des zones de l'image et non pixel à pixel. Cela nécessiterait une classification de l'image qui augmenterait le coût d'implémentation.
- Sinon, pour garder l'étude pixel à pixel, l'évaluation du bruit analogique permettrait de contrôler les distances à prendre en compte qui ne seraient pas concernées par le bruit. Il faudrait dans ce cas différencier le bruit de l'information de l'image, ce qui augmenterait également son coût d'implémentation.

Les informations numériques peuvent encore être visualisées sur un support analogique et donc soumises à ce type de bruit. Néanmoins, les téléviseurs analogiques, par exemple, ont tendance à disparaître du marché. C'est la raison pour laquelle il nous semble plus judicieux d'analyser la robustesse au bruit numérique. De plus, les améliorations proposées pour le bruit analogique étant coûteuses, elles ne sont pas prioritaires.

Le bruit numérique Enfin, nous avons appliqué un bruit numérique, c'est à dire un bruit de compression aux images originale et floutée. Le résultat obtenu est une augmentation du niveau de flou pour l'image nette compressée : en effet, de l'information a été perdue et l'image nous paraît moins nette. En revanche, pour l'image floue compressée, le niveau de flou a diminué : l'image nous paraît moins floue car les artefacts tels que l'effet de bloc recréent de l'information sur les zones lissées de l'image floue. Face au bruit numérique, notre métrique est totalement corrélée à notre perception. Nous avons fait le test pour la métrique de *Marziliano et al* qui donne également ces bons résultats.

Nous venons de valider l'efficacité du BluM au moyen de tests subjectifs et de comparaisons avec d'autres métriques. Dans notre domaine, le BLuM est uniquement estimé sur la composante de luminance. Mais, comme cette métrique est actuellement utilisée dans d'autres domaines, nous avons voulu renforcer notre étude en analysant le rôle de la couleur dans notre discrimination du flou.

4 Le rôle de la couleur dans la perception du flou

Le flou est généré par la suppression des hautes fréquences. Or, la composante de luminance contient plus de hautes fréquences que les composantes de chrominance. L'intuition est donc de penser que l'étude du flou sur la composante de luminance suffit. Nous avons voulu vérifier cela par une étude subjective et une étude objective.

4.1 L'étude subjective

Pour mettre en évidence ce phénomène, nous avons effectué un nouveau test subjectif. Nous commencerons par décrire le protocole de ce test et nous commenterons les résultats obtenus.

4.1.1 Le protocole du test subjectif

Ce test s'est déroulé dans les mêmes conditions matérielles que le premier test subjectif sur le flou. En revanche, les observateurs et les images ont changé. Les observateurs étaient d'autres étudiants du GIPSA-Lab. La série d'image était quant à elle constituée de 27 images plus ou moins floue en couleur et de ces mêmes images sans les composantes de chrominance.

4.1.2 Interprétation des résultats

Nous avons analysé les résultats en regardant pour chaque image la différence entre le MOS obtenu sur les images couleurs et le MOS obtenu sur les images en luminance. Ceci est représenté sur la figure 5.26 où nous avons classé ces images par ordre de notes pour une meilleure visibilité toujours avec l'intervalle de confiance à 95%.

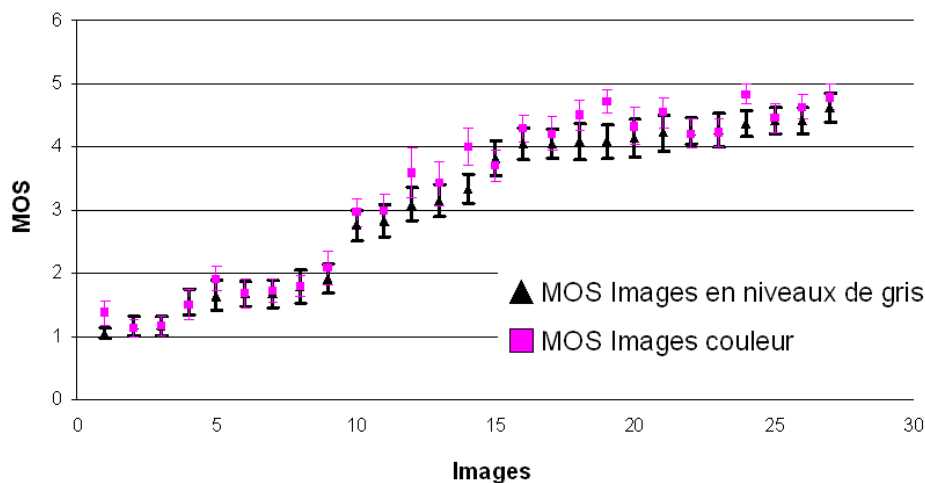


FIG. 5.26 – La perception du flou sur des images avec et sans couleur.

Nous remarquons sur ce graphique que les notes sont proches entre une image couleur et une même image en luminance. L'écart entre ces notes ne fait jamais basculer le jugement d'un palier de qualité à l'autre ce qui montre que la couleur ne joue pas un rôle significatif dans la perception du flou. Enfin, nous remarquons que les MOS obtenus pour les images couleur sont légèrement supérieurs aux MOS obtenus pour les images en luminance et que ce phénomène est accentué pour les notes moyennes de qualité : cela peut s'expliquer par le fait que la couleur offre un

confort visuel qui pousse à mieux apprécier le rendu de l'image et donner une note meilleure. Pour conforter cette théorie, nous avons effectué une étude objective.

4.2 L'étude objective

Nous proposons d'appliquer le principe du BluM mais en tenant compte de la couleur. Pour ce faire, nous avons d'abord dû faire le choix de l'espace couleur.

4.2.1 Le choix de l'espace couleur

Les problèmes liés à l'espace (YCbCr) Nous avons vu dans le premier chapitre que l'espace utilisé pour la vidéo était l'espace (YcbCr). Or, les différents formats 4 :2 :2 ou 4 :2 :0 auxquels nous avons accès ne nous permettent pas d'exploiter les informations de chrominance comme celles de luminance. En effet, sur la figure 5.27, on remarque que les variations d'informations entre pixels voisins sur la chrominance sont différentes car beaucoup plus faibles que les variations d'informations sur la luminance. On ne peut donc pas appliquer tel quel le principe de la métrique de flou sur les composantes de chrominance.

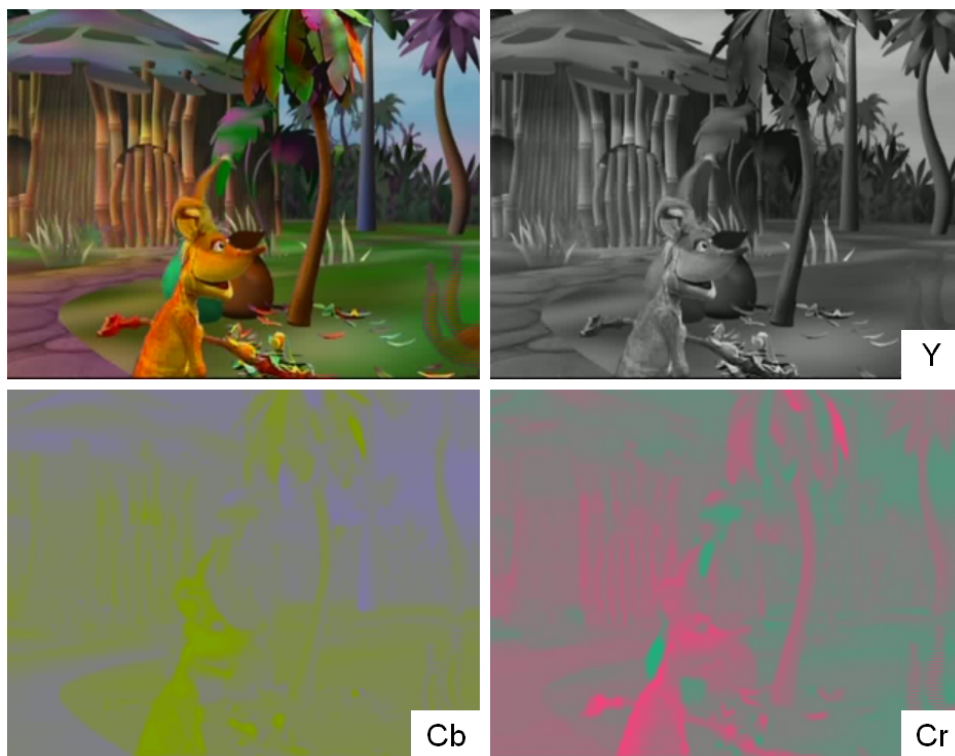


FIG. 5.27 – Une image couleur décomposée en Y, Cb et Cr

De plus, même si on adaptait ce principe pour les composantes Cb et Cr, on aurait une mesure de flou par composante que l'on ne saurait pas exploiter. Ces mêmes problèmes se retrouvent avec l'espace CIE RGB.

L'intérêt de l'espace (L*a*b*) ou CIELAB Nous allons justifier notre choix d'utiliser cet espace en expliquant brièvement comment et pourquoi il a été créé.

En 1942, MacAdam a montré que dans l'espace XYZ (lié à l'espace CIE RGB par une combinaison

linéaire), les distances euclidiennes entre couleurs pouvant tout juste être distinguées dépendent fortement de leur position. La figure 5.28 présente le diagramme de chromaticité et indique à l'aide des ellipses les distances entre couleurs pouvant tout juste être différenciées. Ainsi, un écart couleur dans l'espace XYZ n'aura pas la même importance pour le SVH s'il se produit dans la zone des teintes rouges ou celle des teintes vertes.

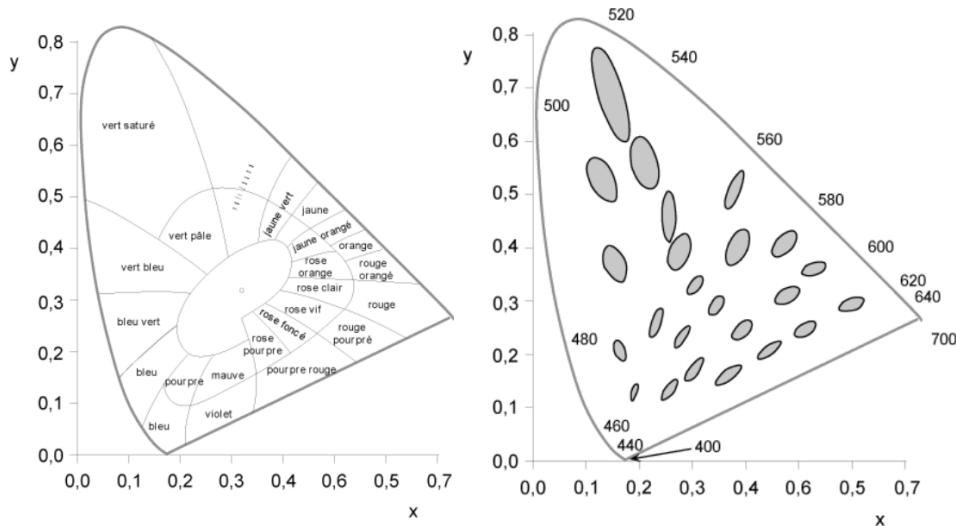


FIG. 5.28 – A droite : diagramme de chromaticité avec position de quelques couleurs. A gauche : Ellipses de MacAdam dans le diagramme de chromaticité

Pour remédier à ce problème, plusieurs systèmes colorimétriques, dont le système CIELAB, sont proposés pour que la distance entre des nuances de couleur proches corresponde à la distance perceptive entre ces couleurs. Les composantes (L^* , a^* , b^*) d'un tel système sont obtenues à partir des composantes (X , Y , Z) par des relations complexes qui rendent difficile son utilisation à faible coût et temps réel.

Dans cet espace, la distance colorimétrique notée ΔE_{ab}^* ou ΔE_{76} est définie comme la distance euclidienne entre 2 couleurs C1 et C2 :

$$\Delta E_{ab}^* = \Delta E_{76} = \sqrt{(L_{C1}^* - L_{C2}^*)^2 + ((a_{C1}^* - a_{C2}^*)^2 + ((b_{C1}^* - b_{C2}^*)^2)} \quad (5.26)$$

En général, on considère qu'un écart colorimétrique inférieur à $\Delta E_{ab}^* = 1$ n'est pas perceptible et on considère comme acceptable un écart dont le ΔE_{ab}^* est inférieur à 4.

Nous avons donc choisi l'espace CIELAB qui est un espace pseudo uniforme, c'est à dire pensé pour qu'il soit possible de calculer un écart couleur entre deux pixels quasiment uniforme quelque soit la teinte.

Remarque : D'autres espaces et mesures de ΔE ont été proposées en 1994 et en 2000. Dans la mesure où notre métrique de flou sur la luminance est obtenue avec des simples différences en valeur absolue, il nous a semblé suffisant d'utiliser l'espace CIELAB pour analyser le rôle de la couleur. Le lecteur intéressé par ces évolutions pourra se référer à l'ouvrage [Trémeau et al., 2004].

4.2.2 Le principe de la métrique de flou en tenant compte de la chrominance

Pour adapter le BluM aux composantes de luminance et de chrominance, nous devons re-définir l'opérateur de distance. Les correspondances entre le BluM en luminance et le BluM en couleur sont récapitulées sur la figure 5.29. Au lieu d'utiliser la composante Y de l'espace YCbCr, on utilise les 3 composantes de l'espace couleur CIELAB. Puis, nous remplaçons le calcul des différences en valeur absolue entre pixels voisins par le calcul des ΔE_{76} . Pour l'étude de la variation, on garde le même principe qui consiste à analyser les différences (ici les ΔE) qui ont diminué entre l'image à estimer et l'image rendu floue. Les formules s'appliquent de la même façon et on obtient une valeur entre 0 et 1.

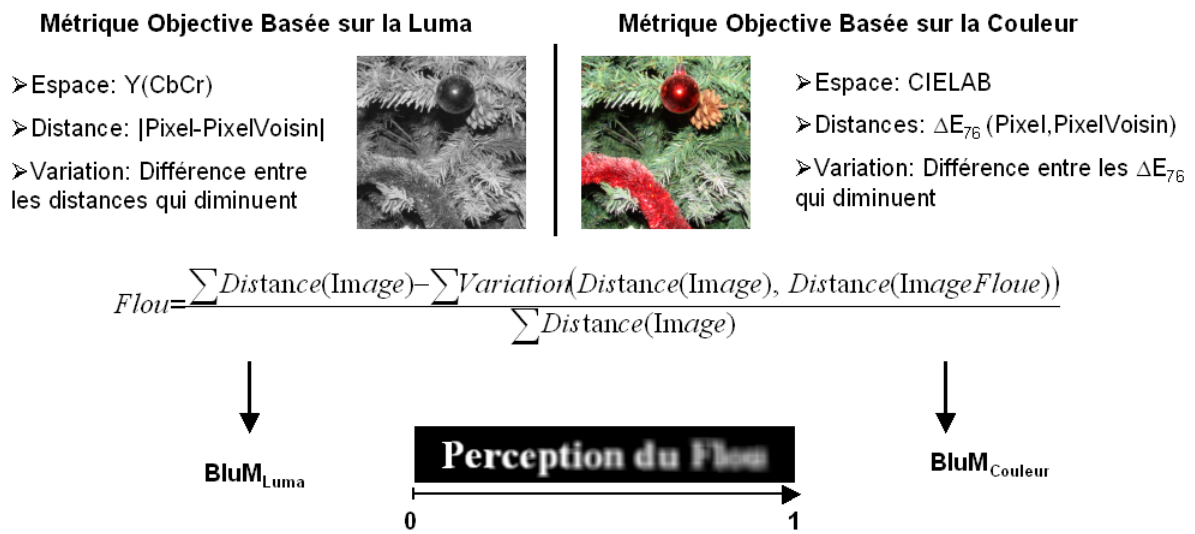


FIG. 5.29 – L'adaptation de la métrique à la couleur.

4.2.3 Les résultats

Nous avons appliqué cette métrique à la même série d'images utilisée pour les tests subjectifs et nous avons confronté les résultats au MOS obtenu lors du premier test subjectif. La figure 5.30 représente cette corrélation. Nous avons effectué la même étude en étudiant la corrélation avec

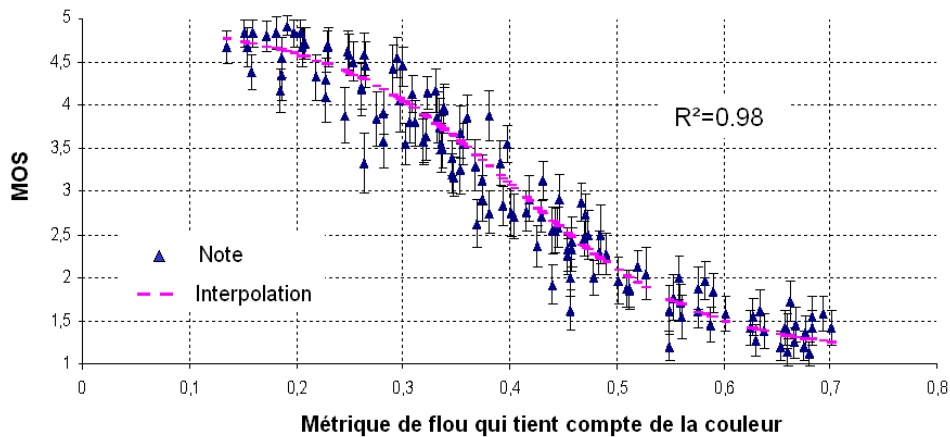


FIG. 5.30 – Corrélation entre le BluM adapté à la couleur et le MOS.

la courbe psychophysique. Nous avons obtenu l'équation suivante qui ressemble à celle obtenue pour la métrique de flou en luminance :

$$A_{Sub} = f(A_{Obj}) = \frac{3,79}{1 + \exp(11,39 \times A_{Obj} - 4,62)} + 1,13 \quad (5.27)$$

La corrélation obtenue est de 0,98 et est meilleure que celle obtenue pour la métrique en luminance. Néanmoins, cette différence n'est pas prépondérante dans la mesure où la première corrélation était déjà bonne.

4.2.4 L'interprétation des résultats

L'étude objective montre que la corrélation est légèrement meilleure mais pas significative lorsqu'on tient compte de la couleur. Ce résultat peut s'expliquer par deux raisons :

- L'écart couleur prend en compte des caractéristiques du SVH alors que la simple différence en valeur absolue des pixels voisins pour la métrique en luminance n'en tient pas compte.
- les tests subjectifs sont effectués sur des images couleurs, ce qui explique que la corrélation soit meilleure pour la métrique de flou en couleur. Pour comparer plus précisément ces corrélations, le test subjectif de la métrique de flou en luminance aurait peut être du être fait sur des images en luminance.

Nous venons de prouver de manière subjective et objective que l'estimation du niveau de flou sur une image est suffisante sur sa composante de luminance. Par ailleurs, cela nous a permis de présenter un principe capable d'adapter une métrique en luminance à une métrique en luminance et chrominance.

Dans ce chapitre, nous avons proposé une nouvelle méthode d'estimation du niveau de flou sur une image. L'originalité de cette méthode est d'utiliser les différentes caractéristiques de la discrimination du flou par le SVH et de les introduire dans un algorithme pour obtenir une métrique sans référence du flou. Ces caractéristiques peuvent être implémentées à faible coût, ce qui rend le BluM exploitable dans notre application.

Puis, nous avons montré l'efficacité de cette métrique en la confrontant à des tests subjectifs réalisés dans les conditions les plus proches possibles de celles recommandées par l'ITU.

Enfin, pour compléter l'analyse de notre perception du flou, nous avons montré de manière objective et subjective que la couleur ne jouait pas un rôle significatif dans la perception du flou mais apportait surtout une sensation de confort visuel pour l'estimation de la qualité d'une image.

Cette métrique de flou nous permet maintenant de compléter l'analyse de la qualité des images compressées et corrigées en la couplant à la métrique de bloc. En effet, sur la figure 5.31, même si les effets de bloc ne sont plus perceptibles, la métrique de flou nous permet dorénavant de dire que la correction utilisée est une mauvaise correction qui a trop lissé l'image et l'a rendue floue.



FIG. 5.31 – La métrique de bloc et la métrique de flou pour évaluer la qualité d'une correction.

Bilan de la seconde partie

Cette seconde partie était consacrée à l'analyse de la qualité des images compressées. Les différentes étapes sont représentées sur la figure 5.32.

Après avoir différencié les notions de qualité et de fidélité, nous avons explicité les deux méthodes les plus fréquemment utilisées pour estimer la qualité :

- Les tests subjectifs nécessitant de la rigueur et du temps mais donnant des résultats fiables.
- Les métriques objectives qui automatisent l'estimation de la qualité au travers d'algorithmes analysant l'image, mais dont la fiabilité dépend de tests subjectifs. Néanmoins, une fois validée, les métriques objectives permettent une analyse rapide de la qualité.

Ce chapitre sur l'estimation de la qualité nous a permis d'aborder ces notions que nous avons utilisées pour commencer à résoudre la problématique.

Un chapitre a donc été entièrement consacré à l'estimation de l'effet de bloc qui est un artefact représentatif de la compression. A partir de la description de cet artefact réalisé dans le chapitre 1, nous avons étudié les métriques proposées dans la littérature que nous pouvions adapter à notre contexte d'étude. En se basant sur ces métriques, nous avons proposé des améliorations et développé le BLE : métrique de l'effet de bloc sans référence. L'efficacité du BLE a été prouvée par des tests subjectifs et par la comparaison avec d'autres métriques. Les limitations évoquées sur le BLE ne sont pas directement liées à notre cadre d'étude. Ce BLE peut donc être utilisé non seulement pour estimer l'effet de bloc dans les images compressées mais aussi pour aider à corriger cet artefact qui est maintenant localisé et mesuré.

Pour compléter l'estimation de la qualité, nous devons être capable d'estimer la gêne occasionnée par les artefacts de correction. Le flou étant le principal artefact de correction, nous avons consacré le chapitre suivant à son évaluation. L'étude de l'état de l'art sur ces métriques nous a montré l'intérêt d'analyser les conséquences du flou sur les contours pour le mesurer. Souhaitant introduire un nouveau principe de détection du flou, nous avons proposé le BLuM qui tient compte de différents phénomènes liés à notre perception du flou. La corrélation entre le BLuM et les tests subjectifs a mis en avant son efficacité.

Ces études nous ont permis d'analyser plus précisément la qualité d'une image compressée ayant subi ou non une correction. L'intérêt est maintenant d'obtenir une méthode de correction capable de donner une valeur de la métrique de bloc et une valeur de la métrique de flou qui correspondent si possible au niveau de qualité "artefact perceptible mais non gênant". Cette étude fera l'objet de la prochaine partie.

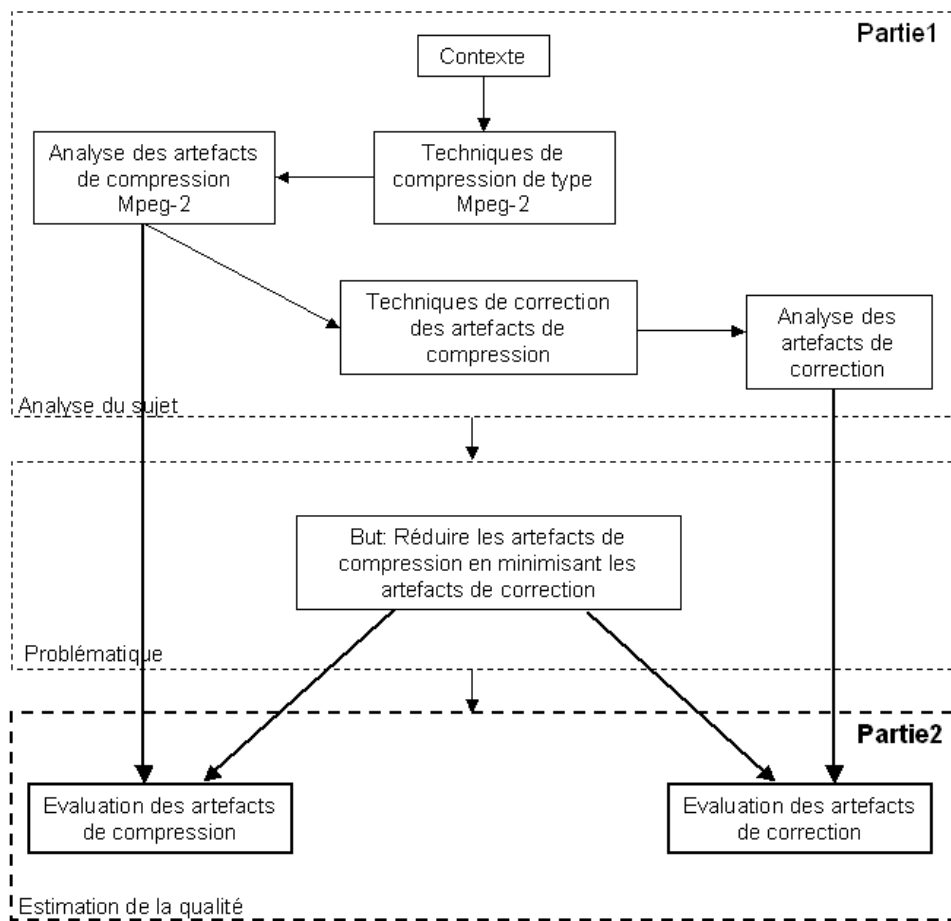


FIG. 5.32 – Bilan de la seconde partie.

Troisième partie

La correction des artefacts

Chapitre 6

Proposition de modules d'amélioration de la qualité des images compressées

Nous avons vu dans le chapitre 2 que la correction de l'effet de bloc et de l'effet de ringing pouvait entraîner d'autres artefacts répertoriés sur la figure 6.1.

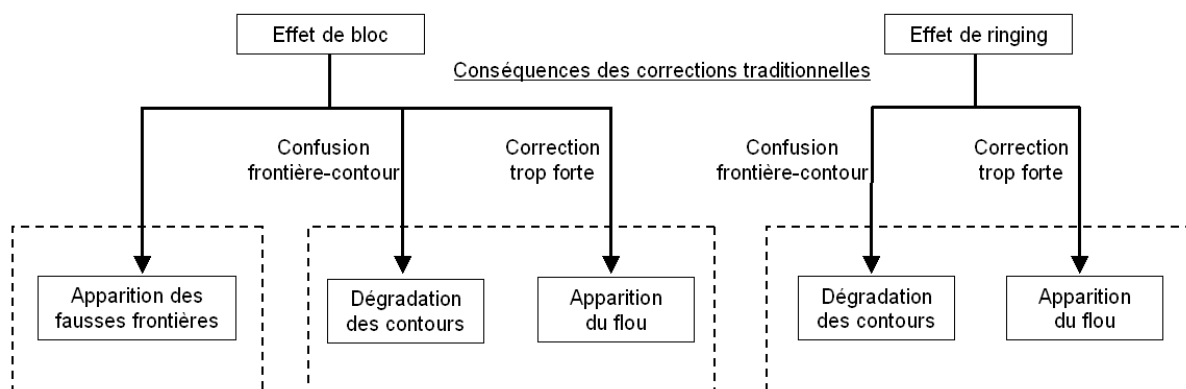


FIG. 6.1 – Rappel sur les causes des artefacts de correction.

Dans ce chapitre, nous proposons trois modules de correction permettant de résoudre ces problèmes.

Dans un premier temps, nous verrons comment éviter la confusion entre les frontières de blocs et les contours qui occasionnent l'apparition du flou après correction.

Puis, nous résoudrons le problème d'apparition des fausses frontières causé par les contraintes d'implémentation en temps réel en adaptant une méthode à ces contraintes.

Enfin, nous contournerons le problème de détection de contours nécessaire à la correction de l'effet de ringing pour proposer un nouveau type de méthode, qui évite la dégradation des contours et l'apparition du flou.

Ces modules sont indépendants les uns des autres et devront répondre aux contraintes d'implémentation. L'efficacité de ces modules sera montrée visuellement sur des images tests et objectivement en utilisant les métriques.

1 Le BLE pour contrôler la correction et éviter toute confusion

Après avoir rappelé la problématique et défini le domaine d'application, nous détaillerons ce module de correction.

1.1 Problématique et domaine d'application

Rappel de la problématique Pour corriger une frontière entre deux blocs, la difficulté n'est pas de trouver le bon filtre mais de ne pas la confondre la frontière avec un contour de l'image. Si ce type d'erreur se produit, le filtrage destiné à supprimer l'information au niveau de la frontière supprimera un morceau du contour. Cela créera une dégradation du contour et un phénomène de flou lorsque cette confusion se reproduit fréquemment dans l'image. Ce type d'erreur est fréquent lorsqu'on n'a pas accès aux paramètres du décodeur et qu'on ne connaît pas le paramètre de quantification qui nous renseigne sur la force de compression du bloc.

Le domaine d'application Lorsqu'une frontière est suffisamment nette pour pouvoir la confondre avec un contour, c'est que la compression est assez élevée. Une correction localisée de type filtrage passe-bas est alors recommandée pour supprimer cette information haute fréquence. Si on applique une correction de ce type, comme le font la majorité des méthodes de suppression de l'effet de bloc, on suppose que seule la frontière de bloc sera supprimée et pas le contenu du bloc. Pour cette raison, nous limitons l'utilisation de ce module aux frontières se situant entre deux blocs homogènes c'est à dire ne contenant pas de variations d'informations.

1.2 La solution proposée

Nous allons expliquer pourquoi et comment nous utilisons le BLE pour rendre plus robustes les corrections dans les zones homogènes.

1.2.1 L'utilisation du BLE

L'utilisation du BLE plutôt que du $V_{I,J}$ On pourrait utiliser la valeur du coefficient $V_{I,J}$ (Chapitre 4 section 2.2.1) estimée pour chaque frontière lors du calcul du BLE. Or, nous avons vu qu'avec ce coefficient, il était encore possible de confondre la frontière et le contour. Cette erreur, acceptable pour les métriques car moyennée avec les valeurs de toutes les autres frontières, entraînera une dégradation dans le cadre d'une correction. Pour cette raison, nous proposons d'utiliser le BLE de l'image entière pour aider à distinguer une frontière de bloc d'un contour.

Le paramètre de décision Si le BLE de l'image est faible, c'est que l'image est faiblement compressée ou que l'effet de bloc est à peine perceptible. Dans ce cas, la probabilité pour que l'information détectée au niveau de la frontière soit un contour est grande, car à faible taux de compression, la frontière de bloc ne pourrait pas être aussi nette.

En revanche, plus le BLE sera élevé, plus le pourcentage que l'information trouvée soit effectivement une frontière sera important. En effet, à fort taux de compression, l'information proche d'un contour est souvent accompagnée de l'effet de ringing ne rendant pas ainsi le contenu des 2 blocs homogène.

C'est pourquoi, nous proposons d'autoriser la correction à partir d'un seuil noté $seuil_{cor}$ et d'adapter la correction en fonction de ce dernier.

La valeur du $Seuil_{cor}$ Pour déterminer la valeur de ce $Seuil_{cor}$, nous utilisons la corrélation effectuée avec les tests subjectifs au chapitre 4 section 3.2 : l'effet de bloc semble non gênant tant que le $BLE < 10$. En dessous de ce seuil $Seuil_{cor} < 10$, il est inutile d'appliquer une correction. Puis, on peut adapter la force de correction en fonction du BLE_t trouvé. En effet, tant que le BLE est perceptible mais non gênant $BLE < 20$, une faible correction suffit. En revanche, pour un BLE très élevé correspondant à des frontières de blocs perceptibles $BLE > 20$, nous proposons d'appliquer une forte correction. En effet, un fort taux de compression fera converger le contenu des blocs vers des blocs homogènes supprimant ainsi l'effet de masquage à l'intérieur des blocs et causant des frontières très perceptibles entre ces blocs homogènes.

1.2.2 L'utilisation du BLE_{t-1} plutôt que du BLE_t

La raison Pour éviter de parcourir toute l'image avant de pouvoir corriger chaque frontière, nous supposons que l'estimation du BLE sur l'image à l'instant $t - 1$ noté BLE_{t-1} est une bonne approximation du BLE sur l'image à l'instant t noté BLE_t .

Les conséquences Le risque de cette méthode est d'avoir un changement important du BLE de l'image t à l'image $t - 1$ et que ce changement nous induise en erreur dans la correction. En effet, lors des changements de scène, le taux de compression augmente et l'effet de bloc est plus visible sur les images I : ce phénomène est représenté sur la figure 6.2 par la valeur du BLE obtenue sur chaque image d'une séquence compressée en MPEG-2. Les deux erreurs possibles sont les suivantes :

- Le BLE_{t-1} est supérieur au BLE_t : le risque est de trop corriger l'image t . Dans ce cas, si on considère l'étude faite par [Tam et al., 1995] mentionnant que le SVH n'est pas sensible si une image dégradée n'intervient que ponctuellement dans une séquence vidéo, cette conséquence n'est pas majeure.
- Le BLE_{t-1} est inférieur au BLE_t : le risque est de ne pas suffisamment corriger l'image t . Dans ce cas, la correction ne dégradera pas plus l'image qu'elle ne l'était.

Les conséquences n'étant pas majeures, nous préférons garder cette estimation sur une image de retard plutôt que de parcourir deux fois la même image.

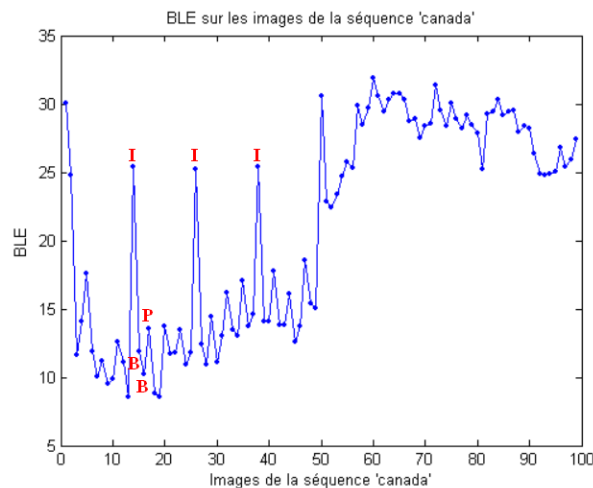


FIG. 6.2 – La distinction des IPB par le BLE sur la séquence *swim*.

1.2.3 La description des corrections

La détection d'une frontière entre deux blocs homogènes Une frontière visible entre deux blocs homogènes doit être détectée. Nous utilisons pour cela le même principe que celui expliqué au chapitre 4 section 2.4.3 : $V_{I,J} = F_{Visible_Homognes}$. Une fois que cette frontière est détectée sur l'image t , nous proposons deux corrections possibles dont la force est directement liée à la valeur du coefficient BLE_{t-1} : si $BLE_{t-1} < 10$, aucune correction n'est effectuée. En revanche, si $BLE_{t-1} > 10$, nous proposons les deux corrections suivantes selon que l'effet de bloc est légèrement gênant ou gênant voir même très gênant.

Le filtrage faible appelé filtrage F1 pour $10 < BLE_t < 20$ Lorsque les frontières de blocs sont légèrement perceptibles, nous préconisons un filtrage faible. Nous avons expliqué dans le chapitre 2 que pour éviter le chevauchement des pixels corrigés et des pixels à traiter, les corrections étaient limitées à $n/2$ rangées de pixels à gauche et $n/2$ à droite pour un couple de blocs de n rangées de pixels. Nous proposons sur cette zone d'appliquer un filtrage passe-bas noté $F1$ dont les caractéristiques sont les suivantes :

- Filtrage à une dimension de noyau $\{1\ 4\ 6\ 4\ 1\}$ sur la zone indiquée sur la figure 6.3. Le choix concernant cette zone est rappelé au chapitre 2 section 3.1.

Son efficacité est montrée sur la figure 6.3. Ce léger filtrage permet de ne pas trop lisser l'information et d'être en accord avec le reste du contenu de l'image représenté par des blocs faiblement quantifiés.

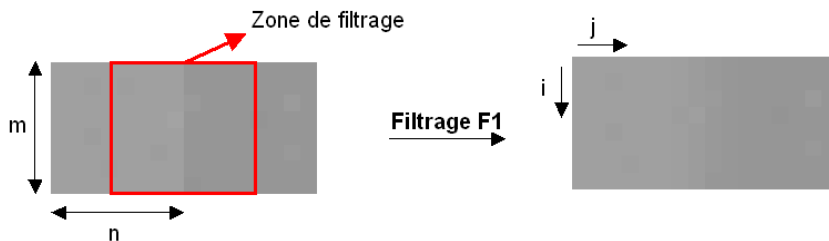


FIG. 6.3 – Résultat du filtrage F1 pour des blocs compressés à taux moyen.

Un filtrage agressif appelé filtrage F2 pour $BLE_t > 20$ Dans le cadre d'une frontière de blocs perceptible et gênante, nous proposons le filtrage F2. Ce filtrage va permettre d'appliquer dans la zone requise un dégradé entre les deux niveaux de gris contenus dans les deux blocs voisins. Ce filtrage de type $F2$ est donné par l'équation suivante :

$$g_1 = B(i, 4) \quad g_2 = B(i, 11) \quad B(i, j + 4) = \frac{((n - 1) - j) \times g_1 + j \times g_2}{n - 1} \quad (6.1)$$

pour i variant de 0 à $(m - 1)$ et j de $n/2$ à $(2n - n/2)$, m et n étant respectivement la hauteur et la longueur d'un bloc.

Le résultat, illustré sur la figure 6.4 sur un exemple de 2 blocs de 8 pixels par 8 pixels montre que la frontière de bloc est plus atténuée avec une correction de type $F2$ qu'une correction de type $F1$.

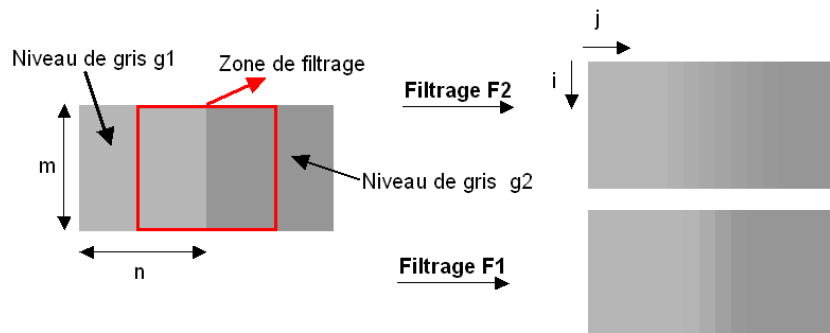


FIG. 6.4 – Résultats des filtres F2 et F1 pour des blocs fortement compressés.

1.2.4 La description du diagramme bloc

Pour répondre aux contraintes d'implémentation, le BLE ne doit pas augmenter la complexité ou le nombre de calculs. Pour cela, nous proposons une analyse représentée sur le diagramme bloc de la figure 6.5. Cette analyse est divisée en trois parties : une étape commune nécessaire à la correction et à l'évaluation du BLE, et deux étapes parallèles :

- L'analyse commune : cette étape analyse la frontière et le contenu des deux blocs et calcule le coefficient $V_{I,J}$. Ce coefficient sera ensuite utilisé pour les deux étapes parallèles.
- Le calcul du BLE : à partir des coefficients $V_{I,J}$, le calcul du BLE sera effectué suivant la méthode décrite dans le chapitre 4 section 2.
- La correction adaptative : parallèlement, si la frontière étudiée correspond à une frontière entre deux blocs homogènes, la valeur du coefficient BLE_{t-1} sera consultée pour adapter le type de correction : pas de correction, filtrage F1, ou filtrage F2.

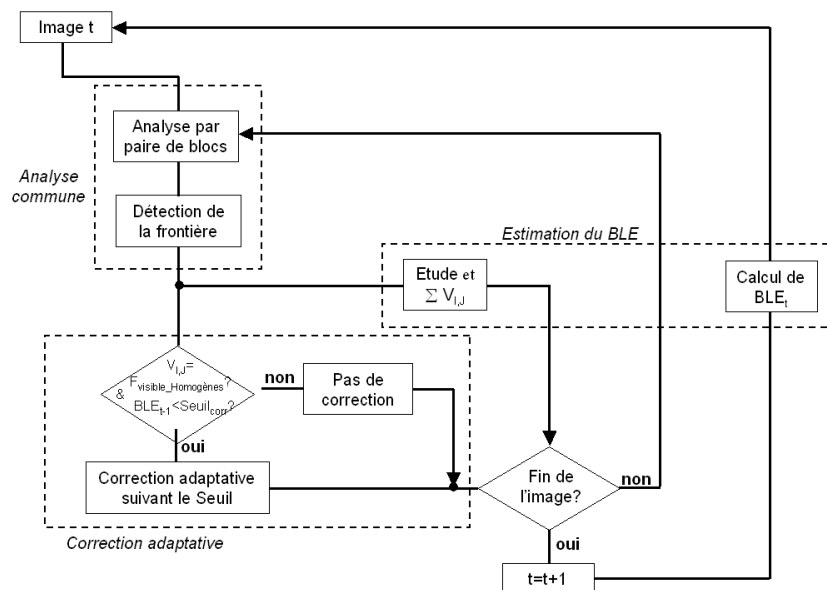


FIG. 6.5 – Diagramme bloc du module d'amélioration intégrant le BLE pour le contrôle de la correction.

1.3 L'efficacité de la méthode

Nous proposons d'illustrer l'efficacité de ce principe qui consiste d'une part à utiliser le BLE pour ne pas confondre une frontière et un contour et d'autre part à adapter la force de correction en fonction de la force du BLE.

1.3.1 La confusion entre frontières de blocs et contours

L'image représentée sur la figure 6.6 est une mire sur laquelle il est fréquent de voir des erreurs de confusions car ces structures linéaires peuvent être confondues avec des frontières si elles sont localisées à des fréquences de 8 pixels.

Nous avons soumis cette image à notre correction de l'effet de bloc guidée par le BLE et à la correction du *MSU Video Group*. Ce groupe propose une méthode de correction de l'effet de bloc qui adapte la taille des filtres en fonction du contenu des blocs, préservant ainsi les détails [MSU,]. Cette correction n'utilise ni l'image de référence, ni les paramètres du décodeur.

Notre BLE ayant estimé le niveau de perception à 6 c'est à dire à un niveau imperceptible, nous n'avons appliqué aucune correction et l'image n'a pas été modifiée. En revanche, la correction du *MSU Video Group* ne faisant aucune estimation de l'effet de bloc et n'ayant pas non plus accès aux paramètres de quantification, dégrade l'image en lissant les contours qu'il a confondu avec les frontières.

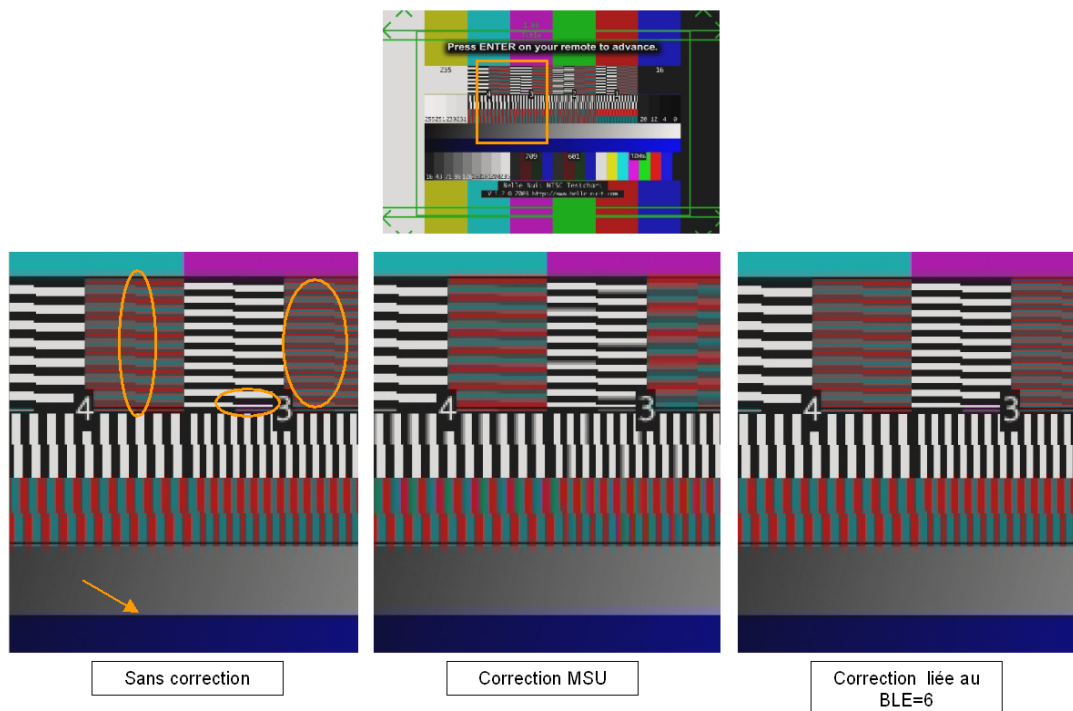


FIG. 6.6 – Les conséquences de la confusion entre les frontières de blocs et les contours et l'intérêt d'intégrer le contrôle du BLE.

1.3.2 L'augmentation du domaine de correction grâce au BLE

De manière complémentaire, nous allons montrer que la force de la correction adaptée au BLE nous permet d'obtenir des corrections plus efficaces. En effet, sur la figure 6.7, nous avons représenté l'image compressée, corrigée par notre principe et corrigée par le *MSU Video Group*. Le BLE trouvé sur cette image étant de 48, un filtrage de type F2, c'est à dire agressif a été appliqué sur les frontières entre blocs homogènes. Cela nous a permis de lisser les frontières beaucoup plus qu'avec une correction classique. En effet la correction effectuée par le *MSU Video Group* fait apparaître des frontières fantômes de part et d'autre de la frontière initiale. On remarque d'ailleurs que la force de correction utilisée est la même que celle de la mire.

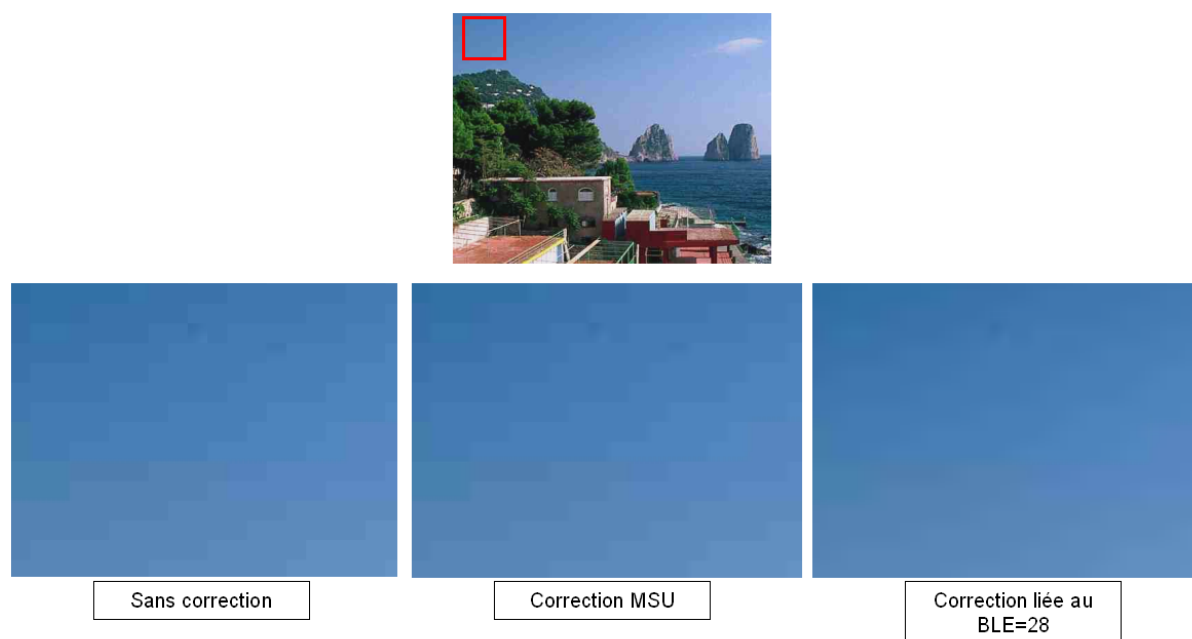


FIG. 6.7 – L'intérêt d'utiliser le BLE pour contrôler la force de filtrage

Grâce à l'utilisation du BLE, nous sommes capables dans des zones avec peu d'informations, de contrôler la force de correction réduisant ainsi les risques de dégradation des contours.

2 La suppression des frontières fantômes

2.1 Rappel de la problématique

La figure 6.8 illustre le problème de chevauchement rencontré avec une correction classique de type classification et filtrage : par exemple, pour analyser la frontière B_{n+1} , nous avons besoin des pixels du bloc Blk_{n+1} qui ont pu être modifiés lors de la correction de la frontière B_n . Pour éviter le problème de chevauchement entre les pixels corrigés et les pixels d'entrée, le nombre de pixels utilisé pour la détection et la correction des frontières est limité à la moitié de la taille de chaque bloc autour de la frontière.

Après une compression par bloc, les zones homogènes sont représentées par une succession de blocs homogènes. Pour une frontière perceptible entre deux blocs homogènes, la principale conséquence de la limitation évoquée précédemment est l'apparition de frontières fantômes G_{n_G} et G_{n_D} à gauche et à droite de la frontière initiale B_n aux positions qui délimitent le début et la fin de la zone de filtrage (figure 6.8). Même si ces frontières fantômes sont moins perceptibles que la frontière initiale, elles rappellent tout de même le phénomène de structure artificielle.

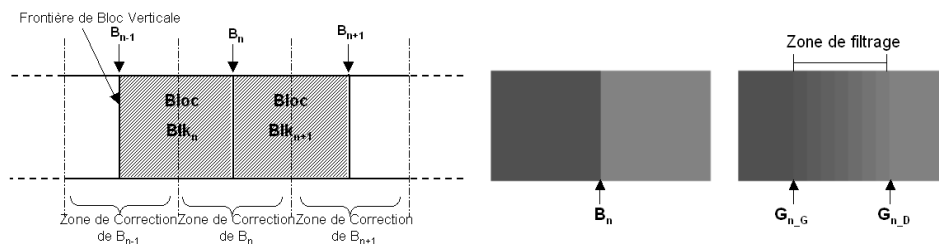


FIG. 6.8 – A gauche : la problématique liée à l'intégration temps réel. A droite : les répercussions avec l'apparition des frontières fantômes

2.2 Domaine d'application

Ce type de problème apparaît lorsqu'une frontière entre deux zones homogènes est tellement visible que la correction sur la zone autorisée par les algorithmes temps réel est insuffisante. Ce cas apparaît dans les zones homogènes lorsque l'information est compressée ou dans les zones avec du détail qui deviennent homogènes après une très forte quantification.

De plus, lorsque certaines zones avec peu de détails sont fortement compressées, nous avons vu dans le chapitre 2 section 3.1 que les valeurs des pixels de zones voisines convergent vers une même valeur et forment des blocs étendus sans frontières intermédiaires. Dans ce cas, il est facile d'imaginer que les frontières fantômes issues de cette correction seront très gênantes car artificielles pour le SVH.

2.3 Les solutions proposées

Pour éviter le problème des frontières fantômes, nous proposons un principe de correction basé sur une observation simple : les frontières fantômes n'existeraient pas si la correction était effectuée sur une zone plus étendue. Cette solution est illustrée sur la figure 6.9 : l'idée est d'utiliser tout l'espace nécessaire pour la correction.

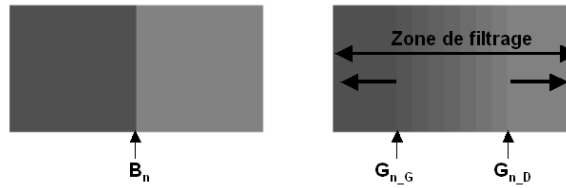


FIG. 6.9 – La solution idéale pour supprimer les frontières fantômes.

Dans le cas des blocs étendus, l'effet de bloc est très prononcé et il est possible qu'une correction sur l'ensemble des deux blocs ne soient probablement pas suffisante. Nous proposons alors d'utiliser l'espace à l'intérieur du bloc étendu pour améliorer la correction par extension de la zone de filtrage. Ce phénomène est représenté sur la figure 6.10.

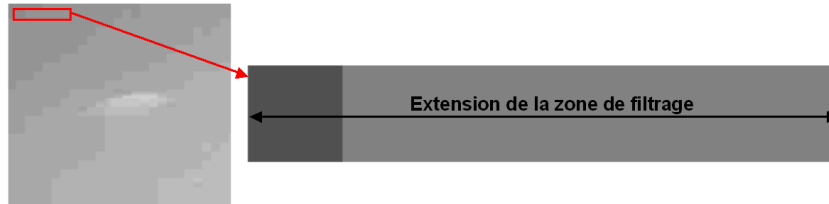


FIG. 6.10 – A gauche : les blocs étendus après une forte compression. A droite : Proposition de correction par extension de la zone de filtrage.

Nous proposons d'expliquer dans un premier temps ces corrections idéales que nous envisageons, puis, nous présenterons les étapes à suivre et parfois les simplifications à faire pour les inclure dans une solution implémentable en temps réel et à faible coût.

2.4 Les méthodes idéales de détection et de correction des frontières de blocs entre zones homogènes

Nous différencierons le cas d'une frontière visible entre deux blocs homogènes du cas d'une frontière visible entre un bloc homogène un bloc étendu.

2.4.1 Cas d'une frontière visible entre deux blocs homogènes

L'étape de détection Pour détecter ce type de frontière, nous utilisons le même principe que celui expliqué dans le chapitre 4 section 2.4.3 formule (4.20).

L'étape de correction Comme nous sommes dans le cas où les deux blocs sont homogènes, nous appliquons le même type de correction que le filtrage F2 à la différence que le dégradé sera étalé sur l'ensemble des 2 blocs pour éviter l'apparition des frontières fantômes. La formule de ce type de correction est donnée par :

$$f_{corr}(i, j) = \frac{[(S_B - 1) - j] \times g_{gauche} + j \times g_{droite}}{(S_B - 1)} \quad (6.2)$$

pour i variant de 0 à $(n - 1)$, j de 0 à $(S_B - 1)$, S_B représentant la largeur en pixels des deux blocs et g_{gauche} et g_{droite} étant respectivement les niveaux de gris à gauche et à droite de la frontière.

Résultat La figure 6.11 illustre le résultat obtenu en utilisant les deux blocs : comparé au résultat obtenu en n'utilisant que la moitié des blocs, nous obtenons un dégradé mieux étalé et la disparition des frontières fantômes.

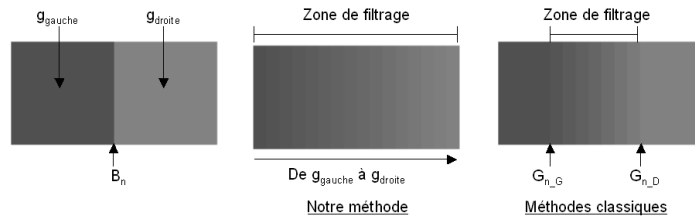


FIG. 6.11 – Résultat obtenu par étalement du dégradé sur les deux blocs et comparaison avec les méthodes classiques.

2.4.2 Cas des blocs étendus

L'étape de détection Dans le cas des blocs étendus, la méthode de détection est la même que celle expliquée dans le chapitre 4 section 2.4.3.

L'étape de correction Une fois que ces frontières invisibles sont détectées, la correction idéale serait d'utiliser tout cet espace pour étaler le dégradé entre g_1 et g_2 en utilisant la formule suivante :

$$f_{corr}(i, j) = \frac{[(S_{BE} - 1) - j] \times g_{gauche} + j \times g_{droite}}{(S_{BE} - 1)} \quad (6.3)$$

pour i variant de 0 à $(n - 1)$ et j de 0 à $(S_{BE} - 1)$ et S_{BE} représente la taille en pixels du bloc étendu.

Résultat La figure 6.12 illustre le résultat obtenu avec cette correction et permet de comparer aux corrections classiques. Nous observons que le dégradé est mieux étalé et les frontières fantômes ont disparu.

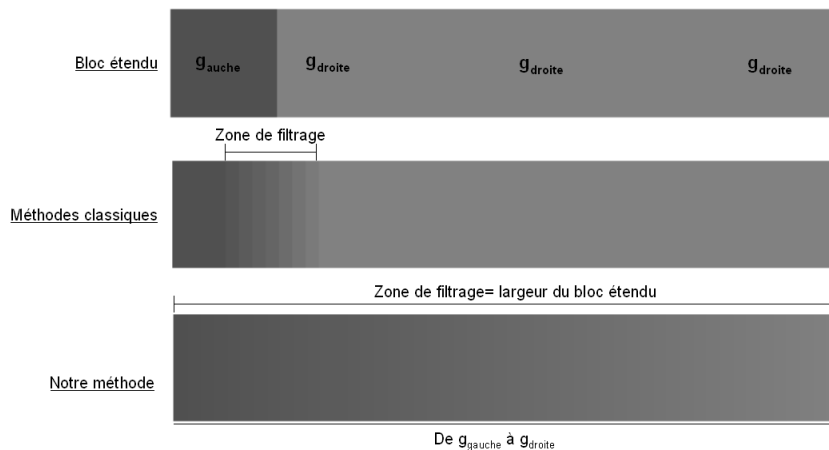


FIG. 6.12 – Résultat obtenu par étalement du dégradé sur le bloc étendu et comparaison avec les méthodes classiques.

2.5 L'intégration dans un algorithme temps réel

La difficulté pour inclure ces corrections est de contourner le problème de chevauchement entre les pixels corrigés et les pixels d'entrée. Nous proposons de décrire les différentes étapes de notre algorithme implémentable en temps réel.

2.5.1 Cas de frontières entre blocs homogènes

Nos propos sont illustrés sur la figure 6.13.

Considérons B_n , une frontière perceptible entre deux blocs homogènes Blk_{n-1} et Blk_n . Dans ce cas, nous appliquons la correction sur tous les pixels en utilisant l'équation (6.5). Nous obtenons deux blocs corrigés Blk_{n-1} et Blk_n . De plus, comme pour le développement du BLE (chapitre 4 section 2.4.4), nous indiquons dans un booléen $Front_{Type}$ en lui attribuant la valeur 1 que si la frontière corrigée était une frontière perceptible entre deux blocs homogènes.

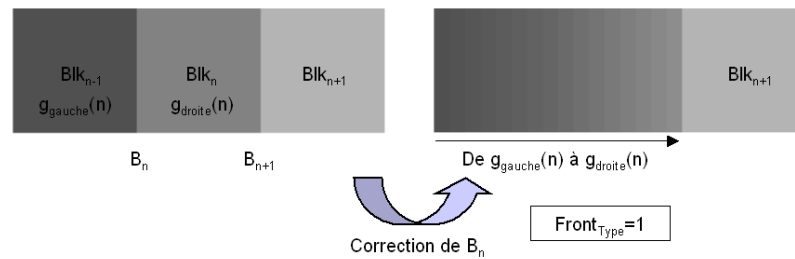


FIG. 6.13 – La première étape de correction avant l'étape de détection.

Pour corriger la prochaine frontière, l'algorithme se déplace d'un bloc et étudie la visibilité de la frontière B_{n+1} entre les blocs Blk_n et Blk_{n+1} . Si Blk_{n+1} est détecté comme un bloc homogène, nous nous reportons au booléen $Front_{Type}$. Si il indique que B_n était une frontière perceptible entre deux blocs homogènes, nous savons qu'avant la dernière correction, le bloc Blk_n était homogène (dans l'état actuel, il contient un dégradé). Nous considérons alors les blocs Blk_n et Blk_{n+1} comme deux blocs homogènes.

Puis, pour étudier la visibilité de la frontière, nous calculons la différence absolue entre la dernière colonne du bloc Blk_n et la première colonne du bloc Blk_{n+1} . L'intérêt de notre correction repose sur le fait que l'application du dégradé n'a pas modifié la valeur de la dernière colonne même si les pixels du bloc Blk_{n+1} ont été modifiés. En effet, avant la correction :

$$Blk_n(i, 7) = g_{droite}$$

et après la correction,

$$Blk_n(i, 7) = f_{corr}(i, 15) = \frac{0 \times g_{gauche} + 15 \times g_{droite}}{15} = g_{droite} \quad (6.4)$$

L'étude de la force de visibilité de la frontière n'est donc pas biaisée par notre précédente correction.

Cette étude nous informe sur l'état de la frontière : perceptible ou non. Si elle ne l'est pas, aucune correction n'est effectuée et le booléen $Front_{Type}$ est remis à 0. En revanche, si la frontière

est perceptible nous appliquons la correction sur l'ensemble des deux blocs en utilisant la formule (6.6). Le résultat obtenu est représenté sur la figure 6.14. On observe une continuité dans le dégradé des niveaux de gris.

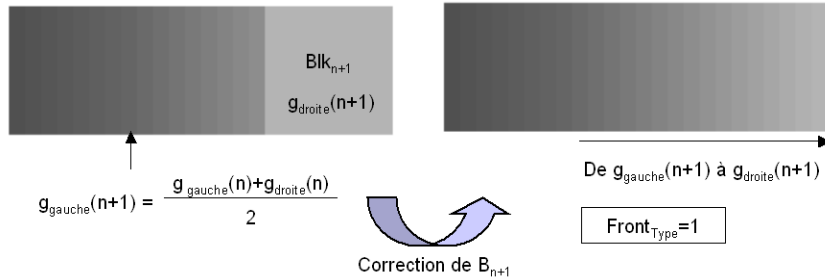


FIG. 6.14 – La correction de la frontière suivante non biaisée par la précédente.

2.5.2 Cas de frontières entre blocs homogènes et blocs étendus

Dans le cas d'un bloc étendu, nous ne pourrions pas effectuer la correction idéale proposée dans le paragraphe (3.4). En effet, pour être implémentable dans un algorithme temps réel, nous pouvons étudier plusieurs blocs consécutifs et compter les frontières invisibles mais nous ne pouvons pas revenir pour corriger la frontière visible. Cela suppose une mémoire plus importante. Nous proposons donc une solution intermédiaire représentée sur la figure 6.15. Cette solution consiste à continuer le dégradé à partir du premier niveau de gris accessible au début de la paire de bloc jusqu'au dernier niveau.

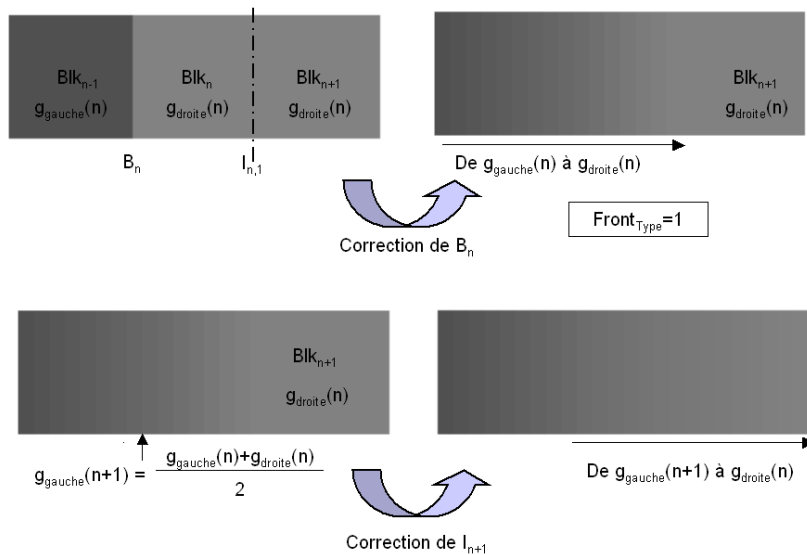


FIG. 6.15 – La correction de la frontière suivante non biaisée par la précédente dans le cas d'un bloc étendu.

La figure 6.16 montre les résultats obtenus pour trois types de corrections dans le cas d'un bloc étendu : une correction classique limitée à la moitié de la taille des blocs, une correction idéale étalant le dégradé sur le bloc étendu et la correction idéale adaptée au temps réel qui continue le dégradé par morceaux. Nous remarquons que même si elle n'est pas aussi bonne que dans le cas idéal, la correction adaptée contribue à une amélioration non négligeable par rapport aux

corrections classiques.

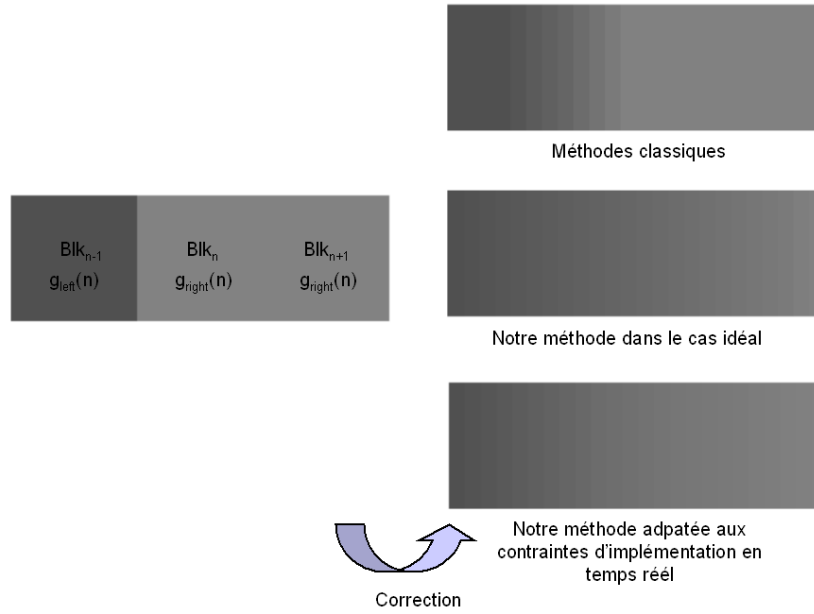


FIG. 6.16 – La comparaison des différentes méthodes de correction en présence d'un bloc étendu.

2.5.3 Remarques et discussion

L'ordre des niveaux de gris Par souci de compréhension, ces principes ont été présentés dans le cas le plus simple où le niveau de gris de chaque bloc suit le même ordre croissant ou décroissant. Si on considère sur une série de trois blocs, g_{gauche} le niveau de gris du premier bloc, g_{milieu} le niveau de gris du second bloc et g_{droite} le niveau de gris du troisième bloc, les formules de correction doivent être adaptées pour le cas particulier suivant : si $g_{gauche} < g_{milieu} > g_{droite}$ ou $g_{gauche} > g_{milieu} < g_{droite}$. En effet, si on applique nos formules, le niveau de gris g_{milieu} ne sera pas représenté. Nous proposons donc les équations générales suivantes :

Si $g_{gauche} \neq g_{milieu}$ et $g_{milieu} \neq g_{droite}$

Pour i variant de 0 à $(n - 1)$ et j de 0 à $(S_B/2 - 1)$:

$$f_{corr}(i, j) = \frac{[(S_B/2 - 1) - j] \times g_{gauche} + j \times g_{milieu}}{(S_B/2 - 1)} \quad (6.5)$$

Pour i variant de 0 à $(n - 1)$ et j de $(S_B)/2$ à S_B :

$$f_{corr}(i, j) = \frac{[(S_B/2 - 1) - j] \times g_{milieu} + j \times g_{droite}}{(S_B/2 - 1)} \quad (6.6)$$

Le résultat obtenu est représenté sur la figure 6.17 et l'amélioration apportée par notre méthode se voit au travers de la continuité de dégradé entre les trois niveaux de gris.

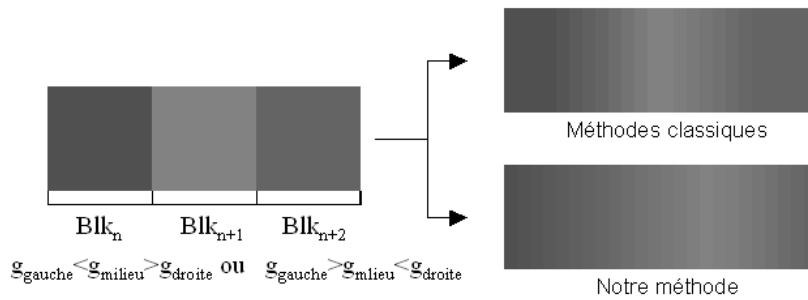


FIG. 6.17 – Correction dans le cas où les niveaux de gris des blocs ne suivent pas un ordre.

La taille des blocs étendus Nous avons défini que la visibilité maximale d'une frontière pouvait atteindre la valeur 32 ce qui correspond à l'écart entre les deux blocs homogènes encadrant la frontière. Dans ce cas, il faut 4 blocs de 8 rangées de pixels pour étaler ces 32 niveaux. Cela montre l'intérêt d'intégrer les blocs étendus. Pour la valeur du booléen $FrontType$, on peut soit décider de limiter la correction à 3 blocs et de le remettre à 0 après la détection d'une frontière invisible, soit décider d'utiliser plus de blocs et de le laisser à 1 tant que les frontières sont soit perceptibles entre 2 zones homogènes, soit invisible mais appartenant à un bloc étendu.

Le cas des frontières horizontales Nous cherchons à proposer une correction implémentable en temps réel. Comme le parcours d'une image se fait informatiquement ligne à ligne, une correction d'un bloc étendu vertical (contenant des frontières visibles et invisibles horizontales) nous obligerait non seulement à gérer un autre booléen, mais surtout à garder en mémoire des blocs de données et à re-parcourir des zones de l'image pour faire la correction. Pour ces raisons, nous décidons de n'appliquer la correction des blocs étendus qu'aux frontières verticales. Cette décision n'a pas de grandes conséquences dans la mesure où les blocs étendus apparaissent sur de grandes zones homogènes telles que le ciel ou la mer et que sur ces zones, la lumière a tendance à être diffusée horizontalement ce qui favorise l'apparition des blocs étendus horizontaux.

Le cas de la chrominance Le phénomène des blocs étendus est causé par la compression d'une variation d'information en luminance. Par exemple, un ciel bleu est généralement représenté par une variation de la luminance pour une même teinte bleue. C'est la compression de la luminance qui va créer les blocs étendus. C'est pourquoi ce phénomène des blocs étendus n'a pas à être appliqué aux composantes de chrominance.

L'utilisation du BLE La correction d'un bloc étendu pouvant être très gênante dans une zone où elle n'a pas lieu d'être, il nous paraît judicieux d'utiliser le BLE donnant une estimation de la compression de l'image avant d'appliquer la correction sur les blocs étendus. Nous décidons ainsi d'activer cette méthode de correction des blocs étendus si le $BLE > 30$.

2.6 Diagramme Bloc de l'algorithme

Nous avons présenté cette correction comme un module pouvant être ajouté à d'autres corrections. En effet, comme le montre le diagramme bloc de l'algorithme sur la figure 7.2, si les blocs de part et d'autre de la frontière contiennent des variations d'informations, rien n'empêche d'attribuer d'autres corrections.

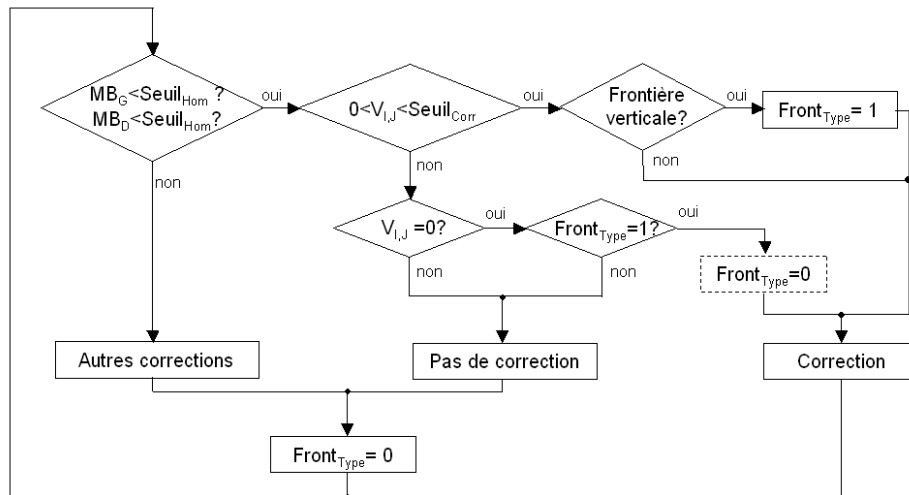


FIG. 6.18 – Diagramme bloc d'une méthode de correction commentant notre module d'amélioration pour éviter le problème des fausses frontières.

2.7 Les résultats

Pour évaluer les résultats, nous avons confronté notre méthode aux méthodes traditionnelles sur des images de différents contenus. Ces tests ont été effectués de manière subjective et objective. Les méthodes traditionnelles ont été implémentées de la même façon que notre méthode mais en limitant la correction des frontières entre zones homogènes à un filtrage de type F2 sur la moitié des blocs.

2.7.1 Les tests visuels

Dans un premier temps, des tests visuels ont permis de comparer l'image originale, l'image compressée, l'image corrigée par des méthodes traditionnelles et l'image corrigée par notre méthode. Deux exemples d'images fortement compressées sont représentées sur les figures 6.19 et 6.20. Sur ces deux images I1 et I2, nous observons que contrairement aux méthodes traditionnelles, notre correction ne fait pas apparaître de frontières fantômes.

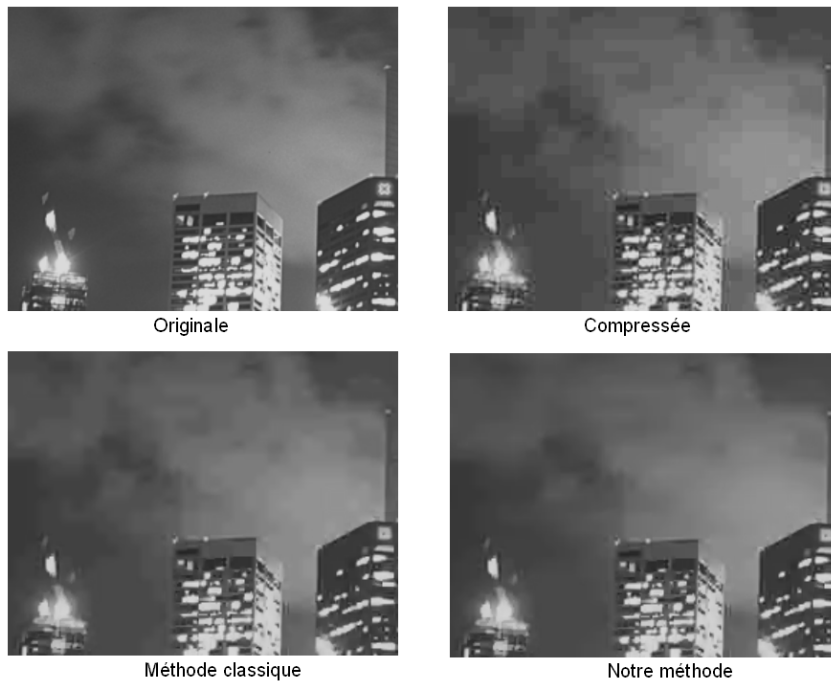


FIG. 6.19 – Image I1 originale, compressée, corrigée avec des méthodes traditionnelles, corrigée avec notre méthode

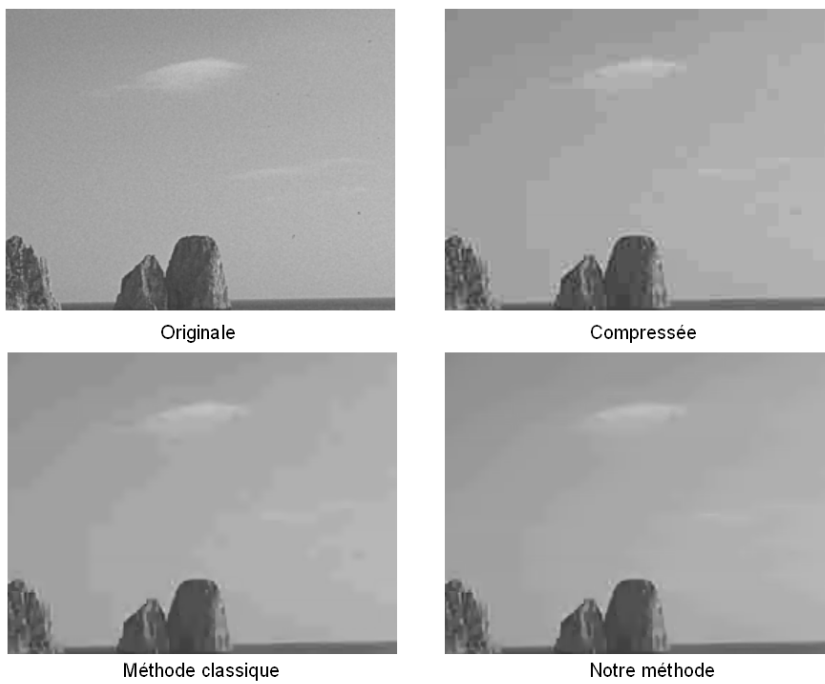


FIG. 6.20 – Image I2 originale, compressée, corrigée avec des méthodes traditionnelles, corrigée avec notre méthode

Notre correction est plus efficace que les autres sur les images fortement compressées et sur les images contenant de petites zones homogènes compressées à taux moyen. Pour les plus faibles taux de compression, nous avons constaté que les 2 méthodes de correction étaient équivalentes. Cela s'explique par le fait qu'il n'y a pas de frontières fantômes.

2.7.2 Les métriques objectives

Pour quantifier cette amélioration, nous avons confronté ces méthodes à des métriques objectives de fiabilité et de qualité. Nous avons choisi comme métriques le PSNR, le SSIM, la métrique de Wang et al [Wang et al., 2002] M1, la métrique de Perra et al [Perra et al., 2005] M2, le BLE et le BluM. Ces mesures correspondent aux images des figures 6.19 et 6.20 et sont représentées dans les tableaux 6.1 et 6.2. Le rappel de l'échelle de ces métriques est représenté sur la figure 6.21.

Très gênant	60	0	1	Très gênant	1	Totale	1
↑				↑		↑	
Effet de bloc	BLE	M1	M2	Flou	BluM	Similarités avec la référence	SSIM
↓				↓		↓	
Imperceptible	0	10	0	Imperceptible	0	Aucune	0

FIG. 6.21 – Rappel des échelles des différentes métriques.

TAB. 6.1 – Récapitulatif des mesures obtenues par différentes métriques sur l'images I1 : A : non compressées, B : compressées, C : corrigées avec des correction classiques, D : corrigées avec notre correction

Image I1	A	B	C	D
BLE	4	53	13	8
M1	9.62	5.12	6.69	7.21
M2	0.22	0.48	0.25	0.26
SSIM	1	0.82	0.83	0.84
MSE	0	199	202	203
BluM	0.17	0.22	0.23	0.23

TAB. 6.2 – Récapitulatif des mesures obtenues par différentes métriques sur l'image I2 : A : non compressées, B : compressées, C : corrigées avec des correction classiques, D : corrigées avec notre correction

Image I2	A	B	C	D
BLE	4	51	12	10
M1	10.02	5.52	7.26	7.32
M2	0.20	0.44	0.27	0.31
SSIM	1	0.78	0.77	0.77
MSE	0	126	129	129
BluM	0.21	0.25	0.26	0.25

Résultats Nous remarquons dans ce tableau que ni le BLE, ni M1, ni M2 ne sont capables d'évaluer la gêne occasionnée par les frontières fantômes et donc l'amélioration de notre correction. Le BLE ne donne pas de différences significative entre les 2 corrections tandis que M1 et M2 évaluent l'effet de bloc comme plus gênant avec notre correction.

La raison est la suivante : tout ce qui est contenu dans le bloc, (même les fausses frontières) est considéré par les métriques comme un effet de masquage qui vient atténuer la visibilité de la frontière. Pour tenir compte de ces frontières, une étude des structures artificielles dans les blocs serait nécessaire mais d'autant plus difficile à distinguer des réelles informations de l'image.

En comparant avec l'image de référence, le SSIM et le MSE ne mettent pas assez l'accent sur la différence au niveau des frontières de blocs.

En revanche, en ce qui concerne le flou, le BluM indique que notre méthode n'augmente pas la valeur du flou par rapport à l'image compressée non corrigée, ce qui signifie que la correction est bien contrôlée.

Nous venons de proposer une méthode capable de supprimer les frontières fantômes tout en respectant les contraintes d'implémentation. Cette technique est validée par des expériences visuelles et le BluM montre que le flou n'est pas augmenté. Une adaptation des métriques de l'effet de bloc pour détecter les frontières fantômes serait nécessaire pour valider de manière objective cette correction.

3 La correction proche des contours et dans les zones texturées

Jusqu'à maintenant, nous avons favorisé les corrections dans les zones homogènes qui se résolvent par des techniques de filtrage simples et non risqués lorsqu'elles sont contrôlées par le BLE. Nous proposons d'étudier le cas où les informations gênantes de compression doivent être supprimées dans des zones plus délicates.

3.1 Domaine de définition

La difficulté de corriger un artefact proche d'une information de l'image est de ne pas dégrader l'information de l'image. Cet artefact à corriger peut être l'effet de ringing proche des contours ou encore une frontière de bloc entre deux blocs contenant des variations d'informations. Ce type de correction est donc préconisé dans les zones où il y a des variations d'information.

3.2 Rappel de la problématique

3.2.1 La problématique liée à l'effet de ringing

La problématique liée à l'effet de ringing expliquée au chapitre 2 section 3.3 montre la difficulté de localiser les contours, de différencier le bruit des contours et de filtrer ce bruit sans endommager les contours. Le problème est le suivant : sur la figure 6.22, un contour fort (en luminosité) et un contour faible sont représentés. Or, par détection de contour par simple gradient, l'information du contour faible est identique à l'information du bruit du contour fort.

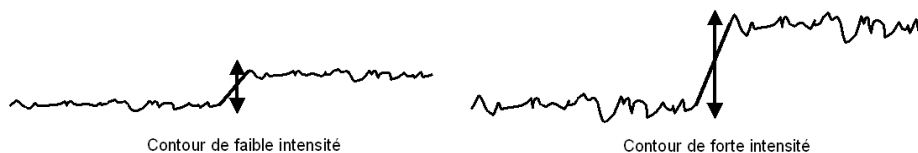


FIG. 6.22 – Schématisation d'un contour de faible intensité et d'un contour de forte intensité entourés de bruit.

Cette information est représentée de manière réelle sur la figure 6.23 qui représente un zoom sur une image compressée et sur l'image gradient correspondante. Les risques encourus sont donc soit de filtrer les contours faibles, soit de ne pas assez filtrer les contours forts.

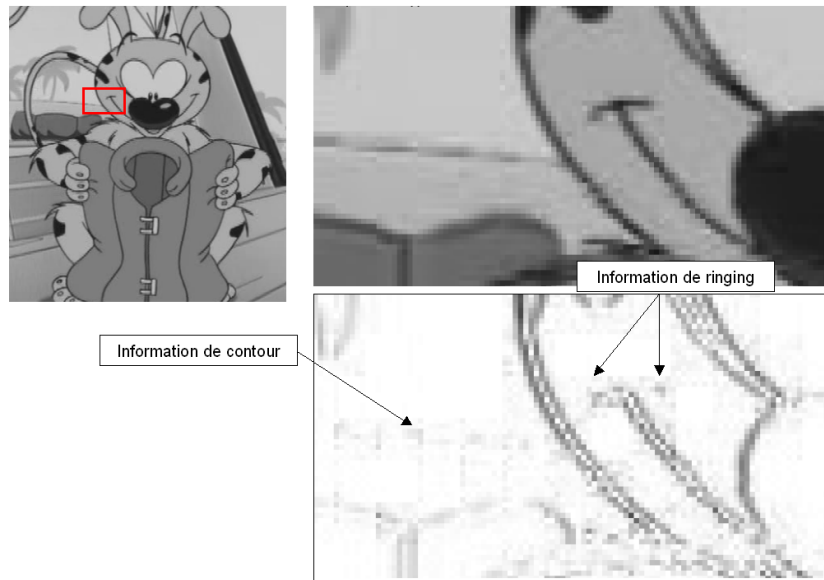


FIG. 6.23 – Contour de faible intensité assimilable à de l'effet de ringing proche d'un contour de forte intensité.

3.2.2 La problématique liée à l'effet de bloc

L'effet de ringing proche des contours est souvent accompagné d'un effet de bloc. En effet, soit :

- Le contour est proche d'une zone homogène. L'effet de ringing se limite au contenu du bloc créant ainsi une discontinuité au niveau de la jonction avec la zone homogène (figure 6.24 à gauche)
- Le contour appartient à deux blocs adjacents. L'effet de ringing proche à chacun des 2 blocs crée une discontinuité de l'information au niveau de la frontière générant ainsi l'apparition de l'effet de bloc (figure 6.25 à gauche).

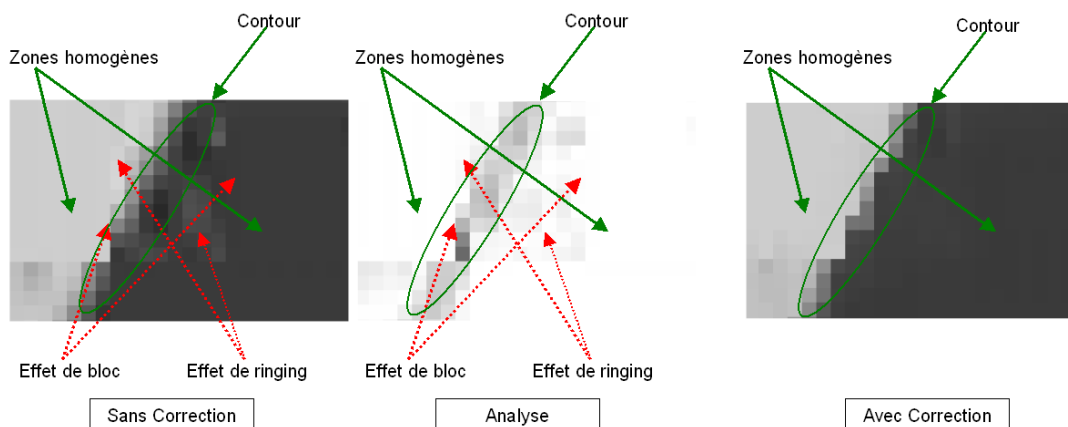


FIG. 6.24 – Effet de bloc entre un bloc contenant un contour et un bloc homogène et effet de ringing le long du contour.

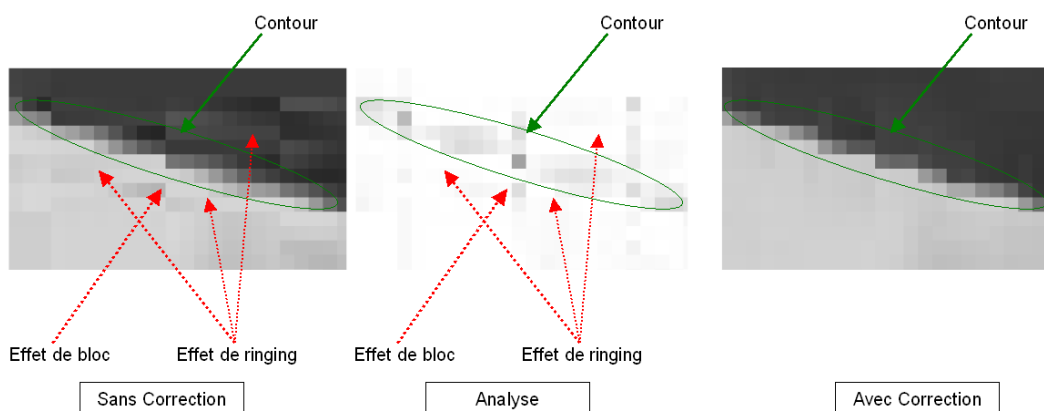


FIG. 6.25 – Effet de bloc entre deux blocs contenant un contour et effet de ringing le long du contour.

Mais la visibilité d'une frontière de bloc n'est pas forcément liée à l'effet de ringing et peut également apparaître entre deux zones texturées comme cela est représenté sur la figure 6.26 à gauche. La difficulté consiste à ne pas apparenter la zone texturée à de l'effet de ringing, auquel cas l'information de l'image sera dégradée par introduction de flou.

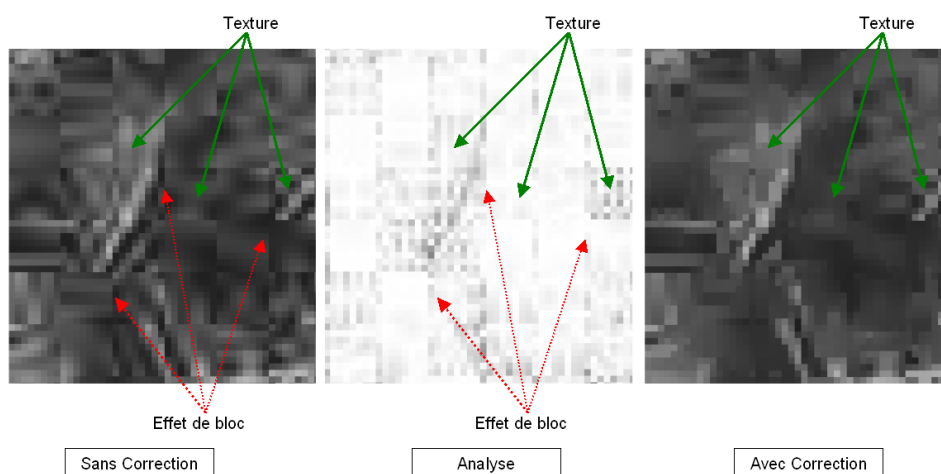


FIG. 6.26 – Effet de bloc entre deux blocs texturés.

3.3 La solution proposée

Nous proposons une méthode générale capable de tenir compte de l'effet de ringing et de l'effet de bloc. Cette méthode doit différencier automatiquement les contours du bruit, et adapter la force du filtrage suivant la force du contour.

3.3.1 Une détection automatique

Nous partons du principe que nous ne devons pas dégrader l'information existante dans l'image compressée sauf si celle-ci est un artefact. Or, dans un bloc, un contour entouré de bruit n'est visible que si l'intensité de ce contour est supérieure à l'intensité du bruit. En effet, sur l'image 6.23, l'effet de ringing du contour prononcé est plus fort que l'effet de ringing du contour moins prononcé. Néanmoins, les deux contours existent et sont mis en évidence par cette image

gradient. La première étape consiste donc à faire un simple calcul de différence entre pixels voisins pour faire ressortir les variations d'informations. Ce phénomène est également représenté par les images du milieu des figures 6.24 et 6.25.

Puis, pour que la correction envisagée ne supprime pas les contours faibles faces aux contours forts, nous normalisons les variations à l'intérieur de chaque bloc. Cela permet de faire ressortir de la même façon, l'information de contour par rapport au bruit.

Les formules suivantes nous permettent d'obtenir cette normalisation.

$$Si \ D(i, j) = |Blk(i, j + 1) - Blk(i, j)|$$

$$norm(i, j) = \frac{seuil_{Max}}{Max - Min} \times (D(i, j) - Min) \quad o \quad Max \neq Min \quad (6.7)$$

$$et \ Max = max(D(i, j)) \quad et \ Min = min(D(i, j))$$

avec i et j parcourant le contenu d'un bloc. Dans le cas où $Max = Min$, cela signifie que les variations sont constantes et que le bruit ne se distingue pas du contour : dans ce cas, aucune correction n'est effectuée.

Toutes les valeurs sont représentées sur une échelle allant de 0 à $seuil_{Max}$. Si le contour est distinguable du bruit, le contour aura des valeurs proches de $seuil_{Max}$ et le bruit aura des valeurs proches de 0. Ainsi, chaque contour, qu'il soit fort ou faible sera évalué de la même manière.

Remarque : Nous venons de présenter ce principe en faisant ressortir les variations verticales mais toutes les directions peuvent être prises en compte. En effet, plus le nombre de directions prises en compte sera important, meilleur sera le résultat.

3.3.2 Un filtrage adaptatif

La détection automatique et la normalisation du contour par rapport au bruit nous permet de ne pas différencier les cas où le contour est faible des cas où le contour est fort. Nous proposons une méthode de filtrage adaptatif basée sur un filtrage appelé filtrage FIV qui signifie Filtrage Inverse aux Variations.

Description du FIV Le principe est le suivant : plus la différence entre pixel voisin est grande, plus l'information est importante et s'apparente à un contour et moins on doit filtrer. En revanche, plus la variation d'information est faible, plus l'information s'apparente à du bruit et plus on filtre. Les formules utilisées pour le filtrage dans la direction verticale sont les suivantes :

$$Si \ S = 1 + T_1 + T_2$$

$$Blk(i, j) = \frac{1}{S} \times (Blk(i, j) + T_1 \times Blk(i, j - 1) + T_2 \times Blk(i, j + 1)) \quad (6.8)$$

$$avec \ T_1 = \frac{1}{D(i, j - 1) + 1} \quad et \ T_2 = \frac{1}{D(i, j) + 1}$$

avec i et j parcourant le contenu d'un bloc. L'ajout du +1 évite la division par zéro.

Les images de droite des figures 6.24 et 6.25 montrent le résultat obtenu sur plusieurs blocs : l'effet de ringing est supprimé et le contour est intact. Nous verrons par la suite comment l'effet de bloc a également été supprimé sur ces images.

3.3.3 L'intégration dans un algorithme faible coût

La figure 6.27 résume les étapes de la méthode proposée pour supprimer l'effet de ringing.

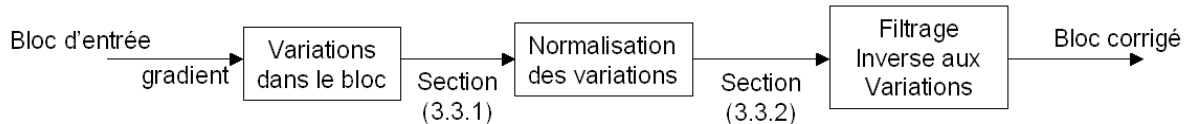


FIG. 6.27 – Diagramme bloc du principe de correction de l'effet de ringing.

Les variations dans le bloc sont facilement mises en évidence par le calcul des différences absolues entre pixels voisins. L'étape de normalisation se fait grâce aux formules (6.9) à (6.11) à condition d'avoir les valeurs minimum et maximum. Or, le calcul des différences pixels à pixels, du minimum et du maximum peut se faire en une seule étape.

La seule difficulté pour intégrer cette solution dans un algorithme faible coût est l'étape de FIV. En effet, les coefficients des filtres sont des coefficients compris entre 0 et 1. Nous proposons pour cela de remplacer ces valeurs par une table de correspondance ou LUT.

3.3.4 L'application de la correction

Cette correction s'applique dans un bloc contenant des variations d'informations. De plus, même si notre correction ne dégrade pas les contours et la texture, nous ne l'appliquerons qu'en présence de bruit de compression. Dans le cas inverse, cette correction pourrait retirer le grain d'une image qui lui donne son aspect naturel. Pour ces raisons, nous appliquons cette correction dans le cas où :

$$Seuil_{cor} > 10 \quad (6.9)$$

3.4 La prise en compte de l'effet de bloc

Nous avons vu que cet effet de ringing faisait apparaître l'effet de bloc à la jonction de blocs adjacents. Cet effet de bloc est légèrement visible car soumis à l'effet de masquage : un faible filtrage est suffisant pour le supprimer. Pour éviter de reparcourir les blocs après le traitement de l'effet de ringing, nous proposons de faire l'analyse de l'effet de ringing et de l'effet de bloc sur deux blocs et de considérer comme nulles les variations au niveau de la frontière. L'avantage est de faire passer la frontière comme une information de bruit et de la filtrer. En effet, sur l'image du milieu de la figure 6.25, on voit clairement ressortir l'information de frontière de bloc. Or avec notre ancien traitement, cette information aurait été perçue comme un contour et n'aurait pas été corrigée. En forçant à 0 ces informations au niveau de la frontière, l'effet de bloc sera considéré comme du ringing ou de l'information non importante et sera filtré.

Dans le cas d'une frontière de blocs entre deux zones texturées, cette méthode permettra de supprimer que la frontière de bloc qui sera vu comme de l'information à filtrer et permettra en plus de préserver la texture qui sera assimilé à des contours. Sur la figure 6.26 au milieu, si on force à 0 l'information de la frontière, cela n'affectera pas la texture.

Le nouveau diagramme bloc associé à cette correction est représenté sur la figure 6.28

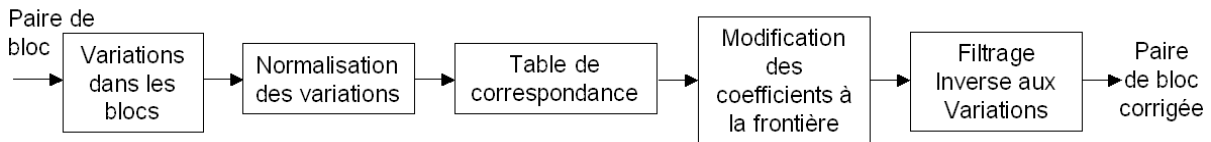


FIG. 6.28 – Diagramme bloc du principe de correction de l'effet de ringing et de l'effet de bloc.

Cette méthode doit être utilisée avec précaution car l'information au niveau de la frontière ne doit pas non plus être confondue avec des contours. L'utilisation du BLE est alors recommandée pour permettre de forcer les valeurs correspondant à la frontière à des valeurs nulles. Nous autorisons cette correction lorsque les frontières de blocs commencent à être perceptibles, c'est à dire lorsque le $BLE > Seuil_{Cor}$ pour rejoindre notre étude précédente.

3.5 Les résultats

Les images de droite des figures 6.24, 6.25 et 6.26 montrent que la méthode proposée permet de supprimer l'effet de ringing et l'effet de bloc en préservant les contours et la texture. La figure 6.29 montre l'efficacité de notre méthode sur des contours de forces différentes.

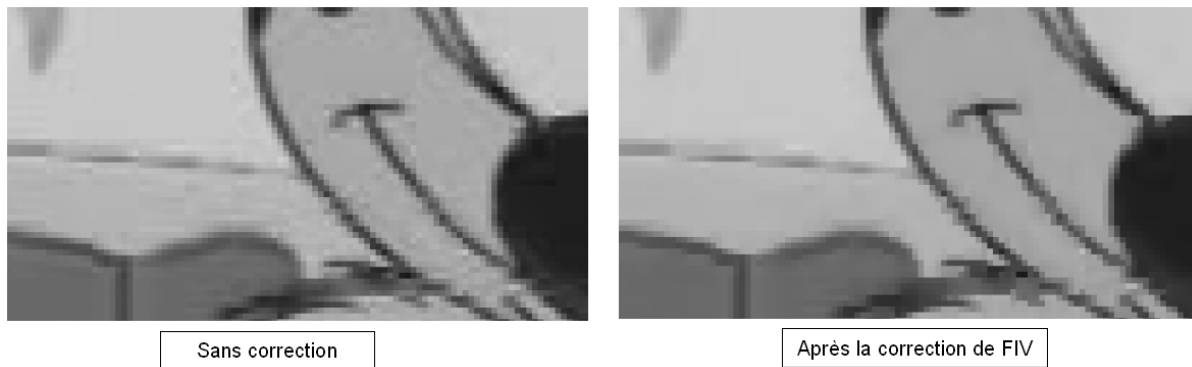


FIG. 6.29 – Zoom d'un résultat de la correction du FIV sur des contours de différentes intensités. Image de gauche compressée, Image de droite corrigée

Remarque : Nous avons utilisé un gradient horizontal et un gradient vertical pour parvenir à ces résultats car, lors d'une correction, l'analyse des frontières verticales et horizontales utilise déjà ce calcul.

Visuellement, nous observons que les contours sont préservés, que le ringing est atténué, et que les frontières entre ces blocs sont supprimées. Si nous appliquons maintenant le *BLE* et le *BluM* sur l'image représentée sur la figure 6.29, nous obtenons :

- Sans la correction de ringing : $BLE = 0.22$ et $BluM = 0.23$
- Avec la correction de ringing : $BLE = 0.15$ et $BluM = 0.22$

Le mesure de flou nous montre que la correction de l'effet de ringing n'a pas introduit de flou, ce qui signifie que les contours n'ont pas été endommagés. La réduction de bruit devrait atténuer la valeur de flou mais cette réduction est compensée par un rehaussement généré par la suppression du bruit autour du contour.

Remarque : L'efficacité de la méthode vient d'être montrée de manière visuelle et avec l'utilisation de nos métriques. Néanmoins, pour renforcer ce résultat, il faudrait appliquer une métrique plus générale tenant compte notamment de l'effet de ringing, comme celle de Caviedes et al [Caviedes et Oberti, 2003].

Nous venons de proposer une méthode capable de supprimer l'effet de ringing et l'effet de bloc lorsqu'ils sont proches d'une zone contenant des contours ou du détail. Des tests visuels, le calcul du BluM et du BLE sur ces images corrigées avec le principe du FIV montre que la correction supprime les artefacts de compression sans faire apparaître d'artefacts de correction.

La figure 7.26 schématise les modules d'améliorations que nous venons de proposer. Ces modules permettent de corriger l'effet de bloc quelque soit l'information contenue dans les blocs, et l'effet de ringing sur n'importe quel type de contour. De plus, l'utilisation du BLE nous permet d'assurer une correction capable de ne pas détériorer les contours. Nous avons également montré que ces algorithmes pouvaient être implémentés avec un algorithme faible coût et temps réel. L'efficacité de ces modules a été prouvée par des tests visuels et par l'utilisation du BLE et du BluM. Ces modules peuvent être greffés à d'autres algorithmes et peuvent être précédés d'un traitement capable de tenir compte des caractéristiques évoluées du SVH.

Nous proposons de réunir ces modules au sein d'un algorithme plus général capable de supprimer

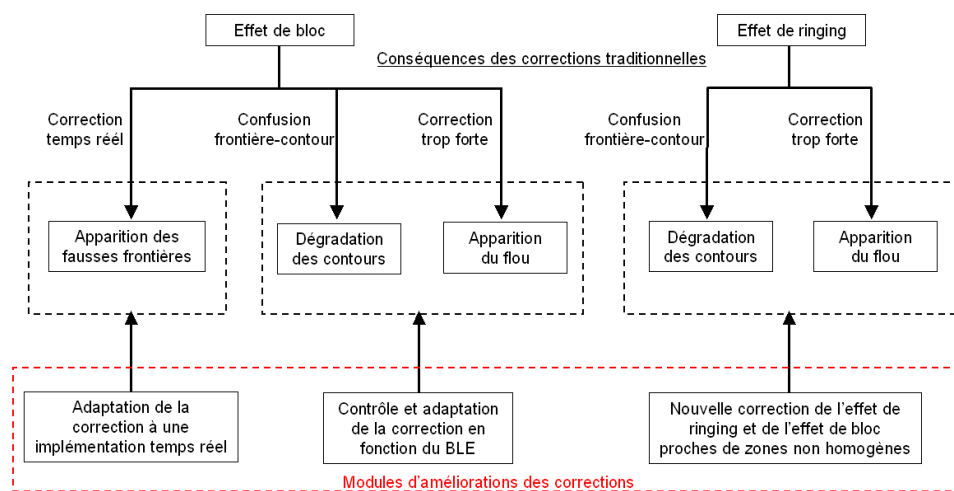


FIG. 6.30 – Modules de suppression de l'effet de bloc et de l'effet de ringing.

l'effet de bloc et l'effet de ringing suite à une compression de type MPEG-2.

Chapitre 7

Proposition et évaluation d'un algorithme de suppression des artefacts de compression MPEG-2

Le but de ce chapitre est d'utiliser nos précédentes analyses sur la qualité et sur la suppression des artefacts pour répondre à la problématique du sujet : comment réduire les artefacts de compression en minimisant les artefacts de correction ? Nous proposons pour cela la description d'un algorithme d'amélioration de la qualité des images compressées appelé SQDSE qui signifie Studio Quality Digital Source Enhancement. La principale application de cette solution est d'intervenir sur les images de diffusion numérique, c'est la raison pour laquelle nous avons ciblé des artefacts de types MPEG-2.

Nous commencerons par décrire la structure de cet algorithme et nous soulignerons les principes utilisés pour le rendre implémentable dans une application faible coût et temps réel. Cette solution reprendra principalement les modules de correction développés dans le chapitre précédent.

Puis, nous analyserons son efficacité en le comparant à une autre solution et en évaluant ses résultats visuellement et objectivement. Nous utiliserons pour cette étape les métriques de qualité. Enfin, nous analyserons les extensions possibles de cet algorithme pour en améliorer son efficacité dans un domaine d'application plus large.

1 Description de l'algorithme

Nous commencerons par définir le domaine d'application du SQDSE.

1.1 Domaine d'application

Le SQDSE traite une séquence, image par image, et suppose que l'on soit capable de lire les pixels d'une image fixe ou d'une séquence ligne par ligne. Les formats d'entrée de ces images sont de types (YCbCr). Les entrées sont la taille de l'image et les valeurs des pixels des lignes correspondantes. En effet, le SQDSE n'a pas besoin de parcourir l'image en entier car c'est une solution qui analyse l'image par blocs et qui restitue cette image par blocs. Nous supposons que ces images sont issues d'un traitement MPEG-2 et que les MBs sont de taille fixe. Or, comme nous souhaitons tenir compte d'un éventuel changement de taille de l'image ou d'un décalage de l'image, le SQDSE a également pour paramètre d'entrée la taille verticale et horizontale d'un MB ainsi qu'un éventuel décalage vertical et/ou horizontal de l'image. Ainsi, le SQDSE sera capable de traiter les 2 images de la figure 7.1 : une image compressée et une image compressée décalée et redimensionnée. Ce type de redimensionnement peut apparaître dans les différents formats d'affichage 4 : 3 ou 16 : 9 et le décalage peut être la conséquence d'intégration de bandes noires pour un recadrage de l'image au centre de l'écran.

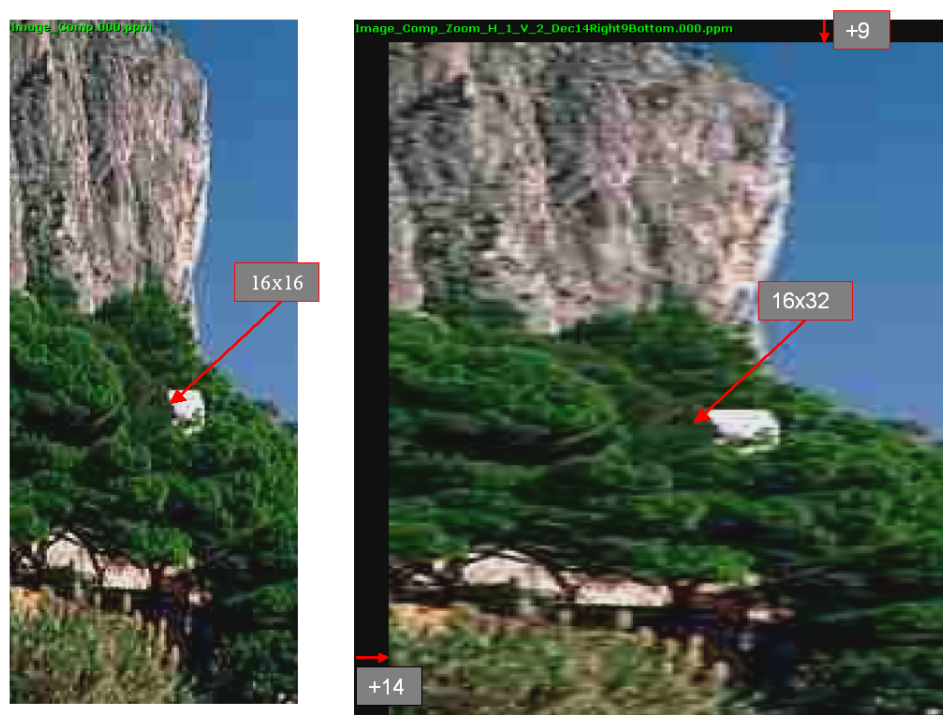


FIG. 7.1 – Image compressée et image compressée redimensionnée et décalée.

1.2 Description du diagramme bloc

Le diagramme bloc du SQDSE est représenté sur la figure 7.2. Nous allons décrire chacune de ces fonctions. Pour simplifier les explications, nos illustrations seront faites sur des MBs de taille 16 pixels par 16 pixels.

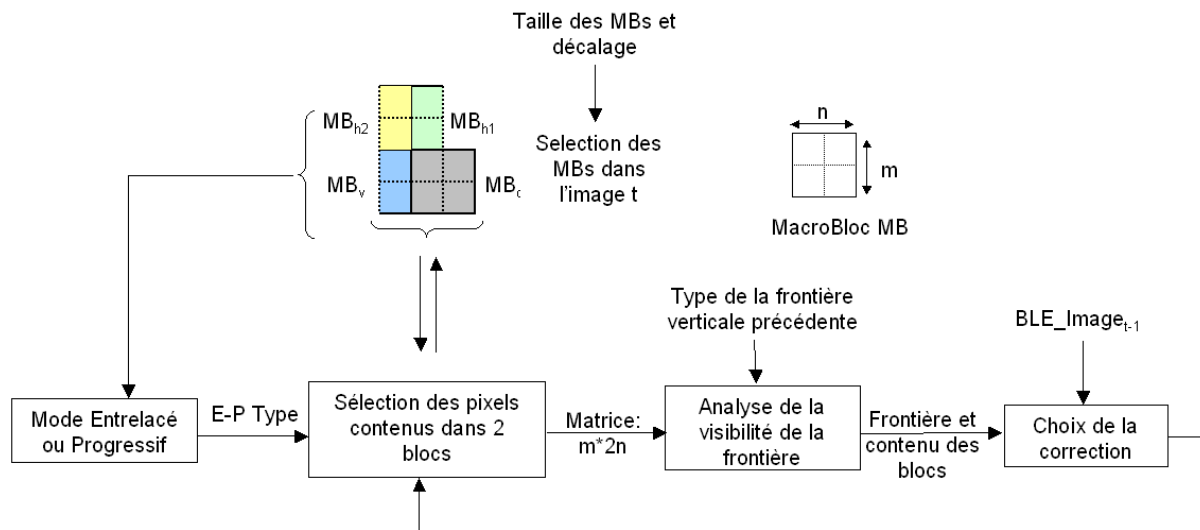


FIG. 7.2 – Diagramme Bloc de l'algorithme SQDSE.

1.2.1 La sélection des MBs

Une fois que nous avons les tailles des MBs et les éventuels décalages de la grille 8x8, nous pouvons lire les pixels de l'image. La lecture de l'image se fait ligne par ligne mais nous avons besoin d'analyser plusieurs lignes pour comprendre le comportement à l'intérieur du MB à traiter. Pour limiter le nombre de pixels à garder en mémoire, nous sélectionnons sur les lignes parcourues en lecture les pixels de 4 MBs : MB_c le parcouru, MB_v le demi MB de gauche, MB_{h1} le demi MB dessus du MB_c et MB_{h2} le demi MB au dessus de MB_v comme cela est représenté sur la figure 7.2. Ce choix précis se justifie par la possibilité d'être en présence de signaux entrelacés et sera expliqué par la suite.

1.2.2 La détection du mode entrelacé ou progressif

Définition Sur la télévision, l'information reçue est des demi images ou demi trames. Une demi trame à l'instant t est composée des lignes paires d'une image et la demi trame à l'instant $t + 1$ est composée des lignes impaires de l'image. Les MBs peuvent avoir été codés en mode entrelacé, c'est à dire qu'ils peuvent contenir les lignes paires de l'instant t et les lignes impaires de l'instant $t + 1$ ou en mode progressif, c'est à dire qu'ils contiennent à la fois les lignes paires et les lignes impaires de l'image à l'instant t .

Dans le cas où le bloc a été codé en mode entrelacé, il faut le désentrelacer pour analyser les pixels dans le bon ordre et ne pas filtrer pixels de l'instant t avec des pixels de l'instant $t + 1$. La figure 7.3 illustre ces 2 cas.

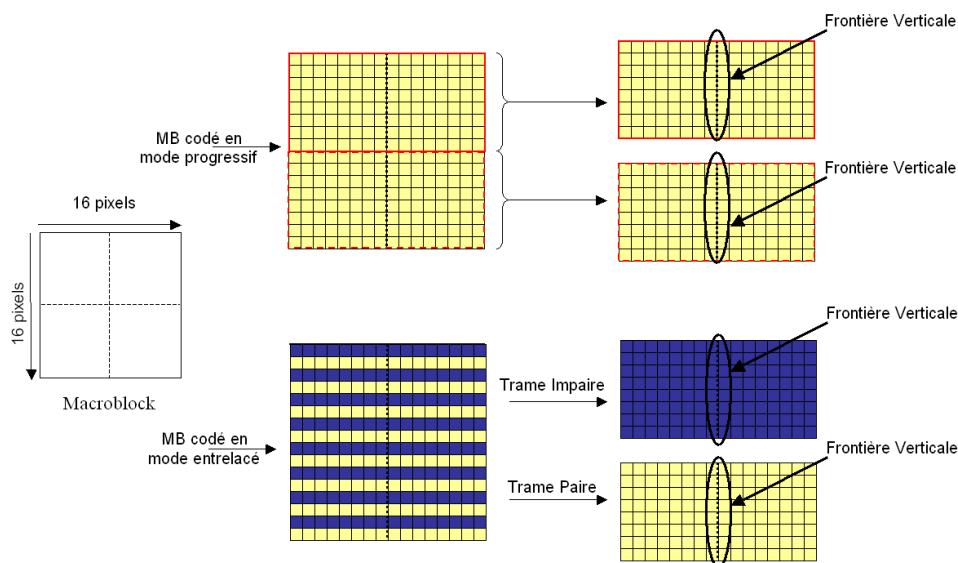


FIG. 7.3 – Cas d’une frontière entre deux blocs dans le mode progressif et dans le mode entrelacé.

Détection Cette information est normalement incluse dans les paramètres fournis par le décodeur. Or, comme nous n’avons pas accès à ces paramètres, nous devons estimer le mode dans lequel chaque MB a été codé. Pour ce faire, une analyse entre la corrélation des pixels ligne à ligne et la corrélation des pixels lignes paires à lignes paires et lignes impaires à lignes impaires est un bon estimateur. En effet, sur la figure 7.4 sont représentés deux MBs réels, l’un encodé en mode entrelacé et l’autre codé en mode progressif. Nous observons que les variations d’information d’une ligne à l’autre sont plus importantes car moins corrélées lorsque le MB a été codé en mode entrelacé.

L’analyse que nous proposons repose sur cette observation et est schématisée sur la figure 7.5 : elle consiste à recréer le MB dans les deux modes et étudier les différences en valeur absolues obtenues entre pixels verticalement voisins. La corrélation la plus importante sera celle obtenue par la plus petite des sommes des différences entre S_{Entr} et S_{Prog} car l’information la plus corrélée donne des variations moins importantes : si $S_{Entr} > S_{Prog}$, on traite le MB en mode progressif et inversement.

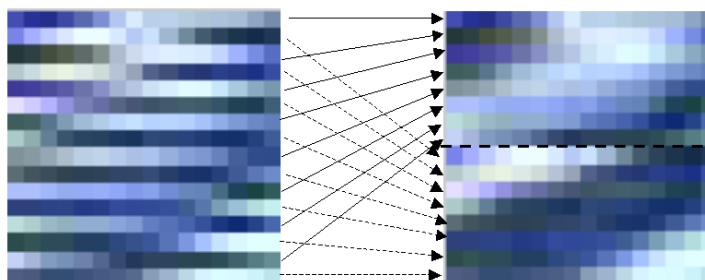


FIG. 7.4 – Cas d’un bloc réel encodé en entrelacé (à gauche) et en progressif(à droite).

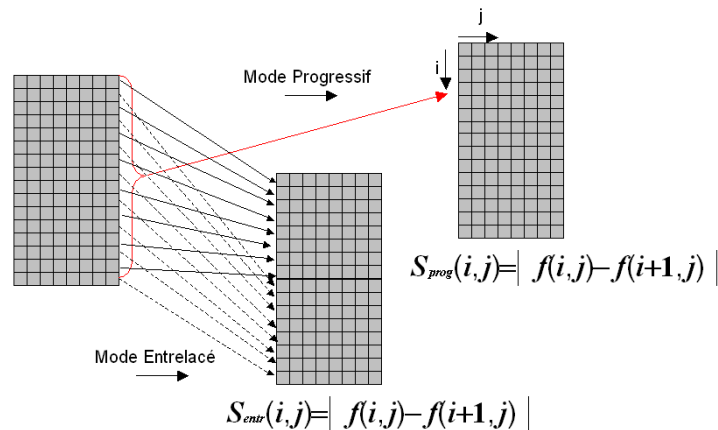


FIG. 7.5 – Détection du mode progressif et du mode entrelacé.

1.2.3 La sélection des pixels

Avec l'information du mode entrelacé ou progressif, nous sommes maintenant capables de sélectionner les pixels à étudier pour l'analyse du contenu des blocs et des frontières. Nous utilisons les pixels issus d'un MB entier et de 3 demis MBs de manière à traiter les 8 frontières représentées sur la figure 7.6. Ces frontières peuvent être verticales, horizontales, entre 2 MBs ou à l'intérieur d'un MB :

- une frontière verticale : 1, 2, 3, 4.
- une frontière horizontale : 5, 6, 7, 8.
- une frontière à l'intérieur d'un MB ou intra MB : 2, 4, 6, 8.
- une frontière entre deux MBs ou inter MB : 1, 3, 5, 7.

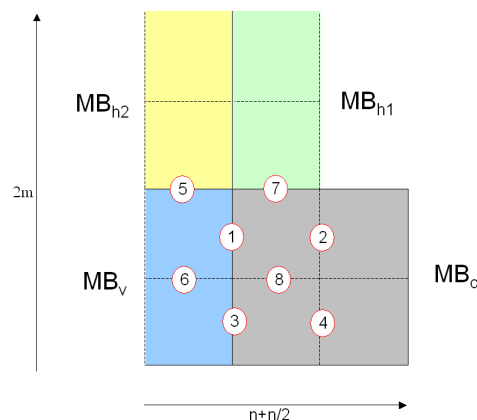


FIG. 7.6 – Les frontières traitées et leur ordre de traitement.

L'ordre représenté par les chiffres est l'ordre de traitement : les frontières verticales puis les frontières horizontales.

Le besoin du MB_v s'explique par la nécessité d'étudier les frontières 1 et 3. La moitié du MB est suffisante car le traitement d'une frontière nécessite juste l'information du bloc de gauche et du bloc de droite.

Le besoin des MB_{h1} et MB_{h2} se justifie par le traitement des frontières inter horizontales 5 et 7. En effet, dans le cas où les frontières ont été traitées dans le mode entrelacé, nous avons besoin des MBs du dessus pour reconstruire l'information. La figure 7.7 représente le cas où les 2 MBs

ont été traité dans le mode entrelacé. Toutes les lignes du MB sont nécessaires pour réorganiser l'information autour des frontières horizontales.

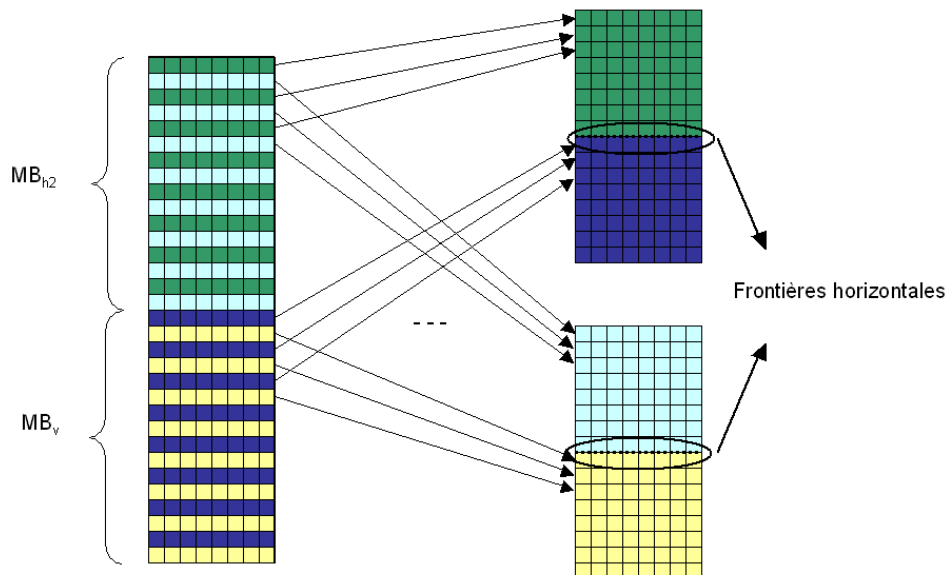


FIG. 7.7 – Cas d'une frontière horizontale entre deux MBs codés en mode entrelacé.

Que la frontière à étudier soit horizontale ou verticale, nous arrangeons nos pixels pour avoir toujours la même matrice à traiter. Dans le cadre de blocs de 8 pixels par 8 pixels, cette matrice a la taille de 8 pixels verticaux par 16 pixels horizontaux ce qui correspond aux deux blocs entre la frontière. Si la frontière est horizontale, une simple transposition de la matrice contenant les deux blocs et la frontière nous permet d'obtenir une matrice de 8 par 16 comme cela est représenté sur la figure 7.8.

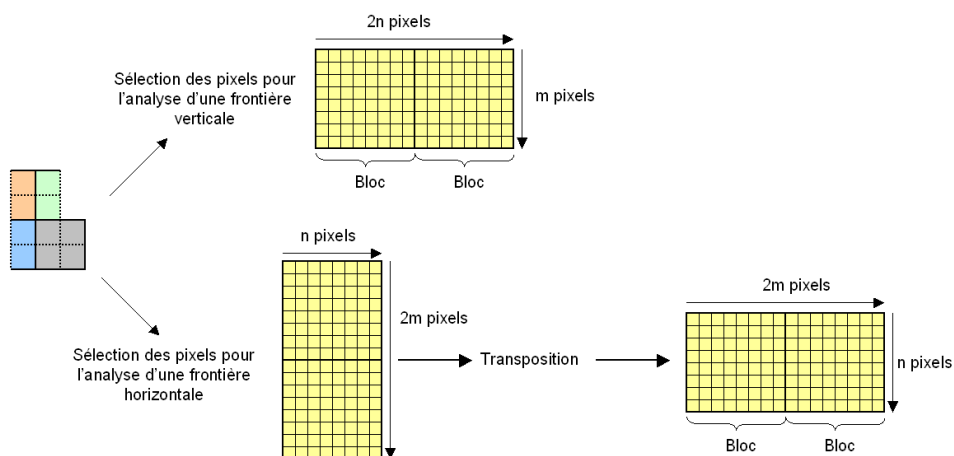


FIG. 7.8 – Cas des frontières horizontales : transposition de la matrice de traitement pour traiter comme une frontière verticale.

Dans le cas où le nombre de lignes est différent du nombre de colonnes, la matrice à traiter aura des tailles différentes mais les traitements seront conservés. De plus, la transposition est une méthode plus simple à implémenter dans un algorithme plutôt que d'avoir à traiter deux cas : frontières verticales et frontières horizontales.

1.2.4 Le besoin en mémoire

Le choix de cette sélection de pixels a été effectué pour minimiser le nombre de lignes à lire et de valeurs de pixels à mémoriser. Cette méthode nécessite la lecture de $2m$ lignes et la mise en mémoire des pixels d'un MB entier et trois demi MBs. Une fois que les 8 frontières correspondant à cette sélection ont été traitées, on décale le traitement d'un MB vers la droite pouvant ainsi traiter 8 nouvelles frontières. Lorsque nous arrivons en fin de lignes, un décalage de m rangées et un retour en début de ligne permettent de traiter les frontières d'une nouvelle rangée de MBs.

1.2.5 Analyse de la visibilité de la frontière et estimation du BLE

Pour chaque frontière à étudier, on sélectionne parmi ces pixels, le contenu des deux blocs de part et d'autre de la frontière et nous étudions la variation d'information au sein de ces blocs et au niveau de la frontière comme cela a été expliqué dans le chapitre 4 section 2.2.1. Si la frontière est considérée comme perceptible, on autorise une correction, sinon, aucune correction n'est effectuée.

Parallèlement, lorsque le MB appartient à la composante de luminance, on stocke la force de la visibilité de la frontière dans une variable temporaire de manière à obtenir le BLE global de l'image une fois que cette dernière sera parcourue. Ce BLE sera utilisé pour guider la correction.

1.2.6 La correction des artefacts sur la composante de luminance

Suite à l'estimation de la visibilité de la frontière et à la valeur du BLE, deux types de corrections peuvent être envisagées :

- La correction de l'effet de bloc par filtrage passe-bas si les blocs contiennent peu de variation d'information
- La correction de l'effet de ringing proche des contours et de l'effet de bloc entre blocs non homogènes.

La correction de l'effet de bloc entre des zones contenant peu de variations d'informations Nous appliquons les méthodes évoquées au chapitre 6 qui utilisent le BLE c'est à dire :

- $BLE < 10$: pas de correction
- $MB_G < Seuil_{Hom}$ et $MB_D < Seuil_{Hom}$ et $10 < BLE < 20$: Correction de type F1
- $MB_G < Seuil_{Hom}$ et $MB_D < Seuil_{Hom}$ et $20 \leq BLE < 30$: Correction de type F2
- $MB_G < Seuil_{Hom}$ et $MB_D < Seuil_{Hom}$ et $BLE \geq 30$: Correction des blocs étendus autorisée

La correction de l'effet de ringing et de l'effet de bloc si les blocs ou au moins l'un deux contient des variations d'information. Dans le cas suivant :

$MB_G > Seuil_{Hom}$ ou $MB_D > Seuil_{Hom}$ et $BLE > 10$:

nous appliquons la correction de l'effet de ringing et de l'effet de bloc expliquée dans le chapitre 6 section 3.

1.2.7 La correction des artefacts sur la chrominance

Une mauvaise correction sur la composante de chrominance peut rapidement causer l'apparition de fausses couleurs lors du regroupement des 3 composantes. De plus, l'information étant sous échantillonnée, notre analyse ne peut être effectuée de la même façon que sur la luminance. En revanche, l'information délivrée par le BLE donne une bonne estimation de la compression en luminance qui est liée à la compression en chrominance. A partir d'un certain seuil de visibilité, à savoir $BLE > 10$, nous autorisons donc, si la frontière est visible une correction légère de type filtrage F1.

Nous venons de décrire le SQDSE qui un algorithme complet capable de détecter et de corriger l'effet de bloc et l'effet de ringing. La section suivante nous permet d'analyser les performances du SQDSE.

2 L'évaluation des performances

Comme précédemment, l'analyse des performances se fera par une étude visuelle des résultats et une étude par mesures objectives. Avant de montrer l'efficacité globale du SQDSE, nous allons détailler l'efficacité du comportement du SQDSE.

2.1 L'efficacité du comportement du SQDSE

Nous avons appliqué le SQDSE sur les images d'une séquence et nous avons soumis les images résultantes au BLE et au BluM. Voyons dans un premier temps l'impact de l'utilisation du BLE sur l'image précédente.

2.1.1 Conséquences de l'estimation du BLE sur l'image précédente

Le graphique de la figure 7.9 représente les résultats du BLE sur 100 images de la séquence *canada* dans les 3 cas suivants :

- non corrigée (bleu et losange),
- corrigée avec le SQDSE en mode temps réel c'est à dire avec le BLE estimé sur l'image précédente (rose carré),
- corrigée avec le SQDSE avec le BLE estimé sur l'image traitée (vert triangle),

Les courbes montrent que l'application du BLE sur l'image à l'instant t permet d'obtenir de meilleurs résultats au niveau des images I pour lesquelles le BLE augmente. En effet, cela permet d'autoriser une force de correction plus forte. Néanmoins, les résultats obtenus dans le cas du SQDSE implémentables en temps réel (avec BLE_{t-1}), sont tout de même corrects avec des valeurs en dessous 15, ce qui correspond en statique, à des effets de blocs perceptibles mais non gênants.

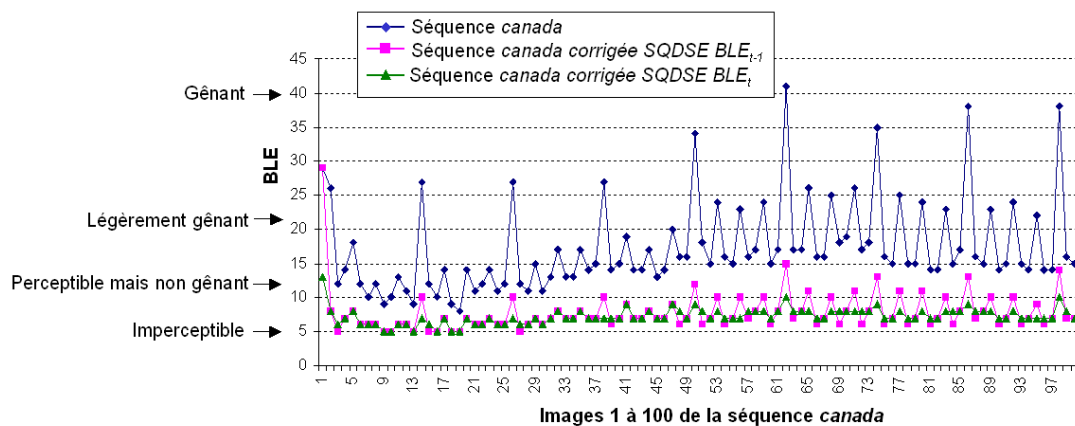


FIG. 7.9 – Conséquences de l'estimation du BLE sur l'image précédente

La figure 7.10 représente la valeur du BluM sur ces mêmes images. Nous remarquons que la courbe (verte et triangle) correspondant à la correction effectuée grâce au BLE estimée à la bonne image donne les points les plus bas ce qui montre la meilleure netteté. Mais la courbe (rose et carré) correspondant à l'estimation du BLE sur l'image précédente donne des valeurs proches qui sont de plus majoritairement plus bas que celles de la séquence non corrigée.

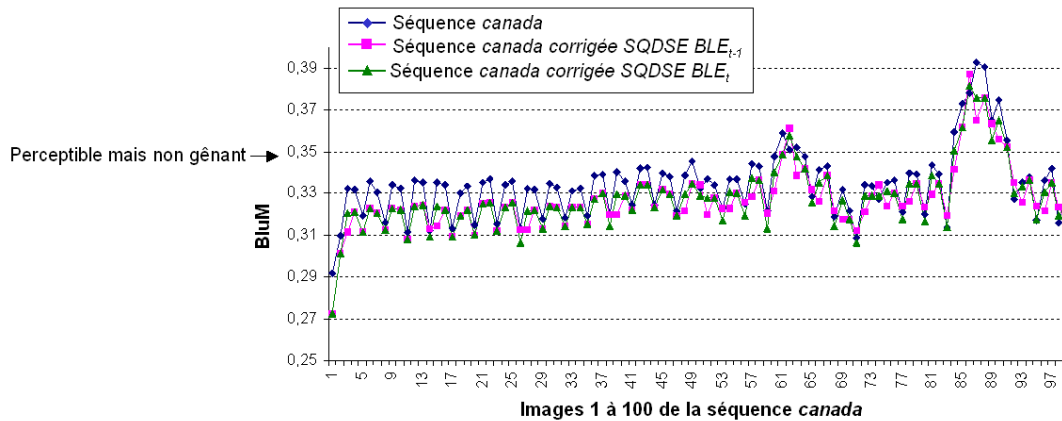


FIG. 7.10 – Conséquences de l'estimation du BLE sur l'image précédente sur le flou

L'estimation du BLE sur l'image précédente n'a donc pas de conséquences importantes sur le manque ou le surplus de correction que cela peut occasionner. Dans la suite du document, la correction *SQDSE* sera donc la correction *SQDSE* avec le *BLE* estimé sur l'image précédente.

2.1.2 Comparaison de l'efficacité de la correction de l'effet de bloc avec une autre méthode

Pour comparer l'efficacité de notre méthode, nous avons soumis les images des séquences testées à la correction proposée par le *MSU Video Group* dont l'efficacité est mise en valeur sur le site proposé à la référence [MSU,]. Nous rappelons que la correction ne concerne que l'effet de bloc. La version utilisée est la version 2.2. Cette méthode peut être utilisée comme filtre avec le logiciel *Virtual Dub* et nous pouvons récupérer les images résultats. Ces images sont ensuite soumises au BLE, ce qui nous permet de comparer les résultats. La figure 7.11 montre les résultats obtenus sur les 100 premières images de la séquence *swim*.

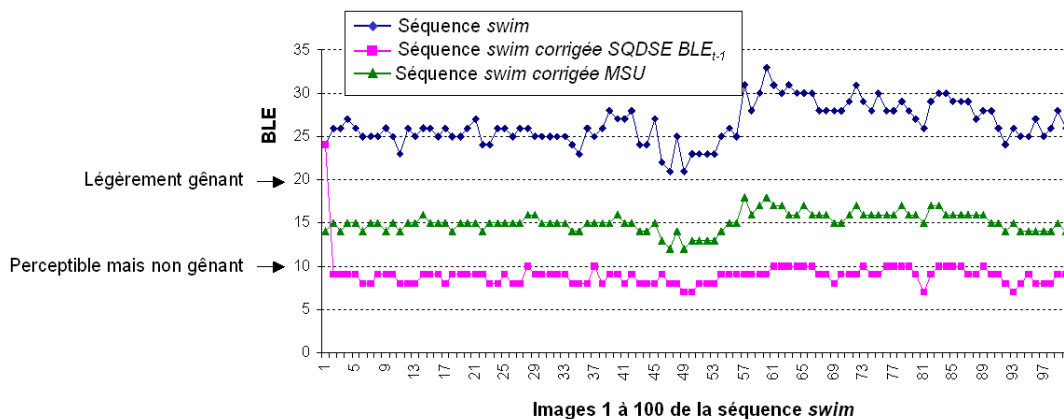


FIG. 7.11 – BLE sur la séquence *swim*, la séquence corrigée par le *MSU Video Group* et la séquence corrigée par le *SQDSE*.

Nous remarquons sur cette figure que la correction du *MSU Video Group* fait bien descendre le BLE à un niveau correspondant à des effets de bloc perceptibles mais non gênants. Le SQDSE supprime plus l'effet de bloc en faisant passer le BLE en dessous de 10, ce qui correspond à un niveau imperceptible.

La figure 7.12 montre les résultats obtenus pour la séquence *canada*. Nous observons que les résultats obtenus par la correction *MSU Video Group* sont moins réguliers que pour la séquence *swim*. En effet, cette séquence contient beaucoup de mouvements qui sont codés par des MBs en mode entrelacé. Ce phénomène non pris en compte par la correction du *MSU Video Group* gêne la détection des frontières et donc sa correction.

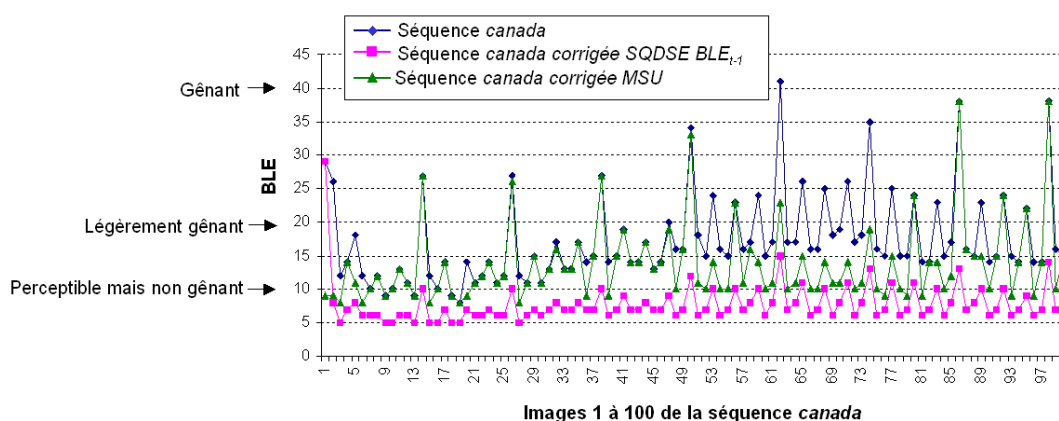


FIG. 7.12 – BLE sur la séquence *canada*, la séquence corrigée par le *MSU Video Group* et la séquence corrigée par le SQDSE.

La figure 7.13 montre le résultat obtenu par le *MSU Video Group* et le SQDSE. On voit que le *MSU Video Group* ne fait aucune correction tandis que le SQDSE supprime les frontières sans pour autant lisser l'information de peigne qui traduit l'entrelacement. Ceci permettra lors de la lecture des lignes paires puis impaires de reconstruire la bonne information sans effet de bloc. De plus, on remarque que même avec la prévision sur l'image précédente, le BLE fait globalement une bonne correction car l'effet de bloc est rendu imperceptible ou dans les pires cas perceptible mais non gênants.

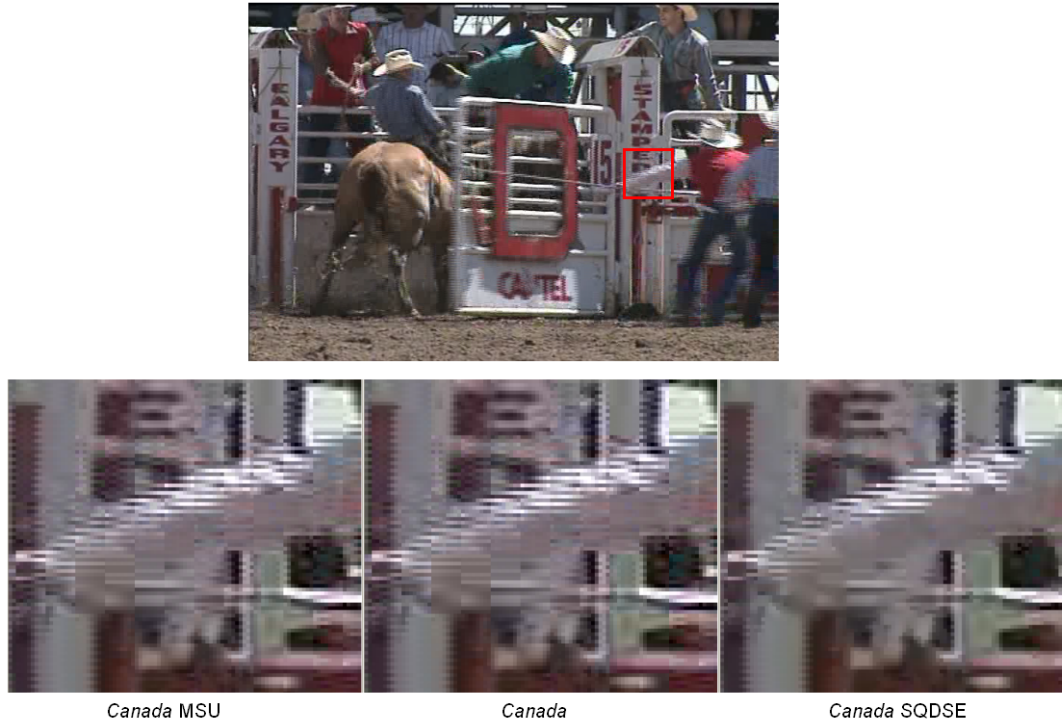


FIG. 7.13 – Effet des corrections *MSU Video Group* et *SQDSE* sur des MBs codés en entrelacé.

2.1.3 Impact des corrections sur la perception du flou

Nous venons de voir que le *SQDSE* diminuait considérablement le BLE sur chaque image d'une séquence. Voyons si cette diminution n'est pas obtenue au détriment de l'apparition de flou sur ces images. Pour ce faire, nous avons soumis ces mêmes séquences au *BluM* pour estimer le niveau de flou. Les résultats obtenus sont représentés sur la figure 7.14.

En observant l'échelle utilisée, la première remarque est que ni la correction du *MSU Video Group*, ni le *SQDSE* n'ont un impact important sur la perception du flou car les valeurs obtenues ne traduisent pas un changement significatif. En revanche, en observant plus précisément, on remarque que la correction du *MSU Video Group* a plus tendance à faire augmenter le flou, contrairement au *SQDSE* qui a plus tendance à diminuer le niveau de flou. Cela s'explique par les précautions prises pour ne pas dégrader l'information de l'image et ne pas la confondre avec l'information de compression.

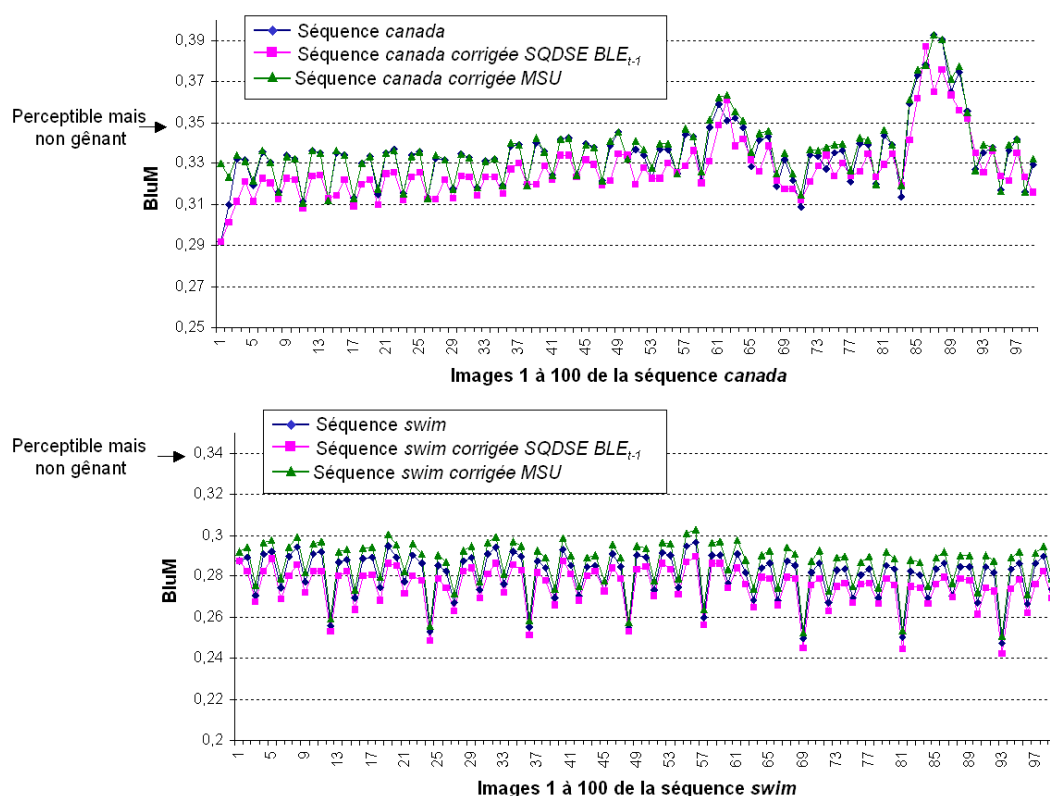


FIG. 7.14 – Mesure du flou sur les séquences compressées et corrigées par le *MSU Video Group* et le SQDSE.

2.1.4 Sur des séquences décalées ou redimensionnées

Nous avons expliqué que le BLE comme le SQDSE pouvait s'adapter à des tailles de MBs différentes dans la mesure où cette information lui est donnée. Par défaut, les MBs sont de taille 16 pixels par 16 pixels. Néanmoins, le SQDSE est conçu et implémenté de manière à pouvoir traiter tout type de taille de MBs ainsi que d'éventuels décalages horizontaux et ou verticaux. La figure 7.15 montre le résultat du SQDSE après un décalage. Le BLE obtenu est quasiment identique avant et après décalage. La légère différence dans la valeur du BLE est due à la différence de contenu créée par le décalage.

La figure 7.16 montre le résultat obtenu après un décalage et un redimensionnement. Le BLE obtenu est meilleur après ce changement car le zoom de type bicubique effectué sous *Matlab* a lissé les frontières ce qui les rend plus difficile à traiter (avant le redimensionnement, $BluM = 0.21$ et après $BluM = 0.32$). Dans ce cas, l'étape de détection ne détecte pas une frontière nette et n'approprie pas une forte correction.



FIG. 7.15 – L'efficacité du SVDSE après un décalage de l'image.



FIG. 7.16 – L'efficacité du SVDSE après un décalage et un changement de dimension de l'image.

2.1.5 La stabilité en temporel

Les résultats illustrés dans les paragraphes précédents montrent que les courbes du BLE et du BluM des images corrigées suivent les courbes des images compressées. Cela signifie que notre correction est stable et donc qu'elle n'apporte pas de dégradations lorsqu'elle est visualisée temporellement.

2.1.6 L'effet de moustique

Nous avons vu que l'effet de moustique se manifestait au niveau des contours car le bruit autour de ces contours variait d'une image à une autre. Le SQDSE corrigeant l'effet de ringing autour des contours, l'effet de moustique doit être atténué par la correction. Sur la figure 7.17, nous avons représenté une image contenant du ringing et de l'effet de bloc avant et après SQDSE. En dessous de ces deux images, nous avons mis en évidence les variations de cette image à l'instant t et à l'instant $t + 1$: plus les variations sont importantes, plus le pixel est foncé.

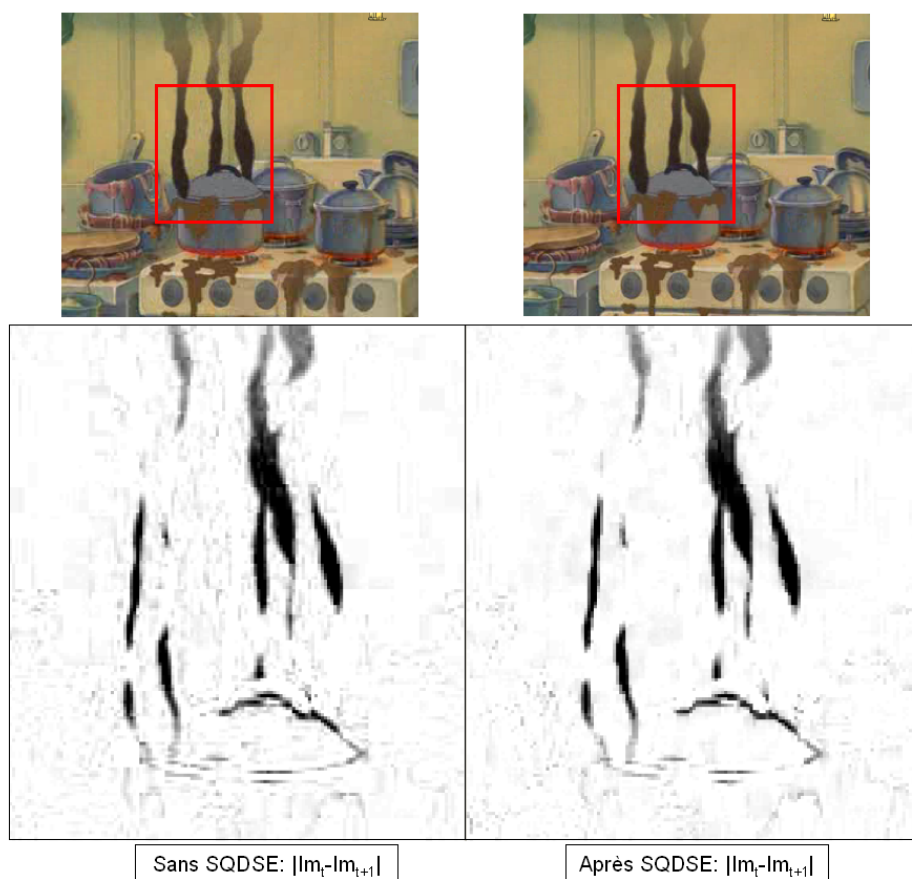


FIG. 7.17 – Impact du SQDSE sur l'effet de moustique : à gauche : image différence entre 2 images consécutives sans correction ; à droite : image différence entre 2 images consécutives avec correction.

Résultat Si on compare les résultats avec et sans SQDSE, on remarque que les variations proches des contours ont diminué. Cela montre que la correction de l'effet de ringing réduit l'effet de moustique car atténue les variations temporelles.

2.2 L'efficacité globale du SQDSE

Nous venons de montrer quelques cas particuliers pour mettre en avant l'efficacité du SQDSE. Cette validation a été effectuée sur 40 séquences de contenus différents et plus ou moins compressées. Nous avons observé les courbes de l'évolution du BLE et du BluM au cours de ces séquences. Pour simplifier la représentation, le graphique de la figure 7.18 ne représente que le BLE et BluM moyen obtenus en moyennant les valeurs obtenues sur l'ensemble des images d'une même séquence : après les corrections du SQDSE, les valeurs moyennes obtenues sont majoritairement en dessous de 10. Cela ne prouve pas que toutes les images sont en dessous de 10 (comme cela a été montré sur la figure 7.12 mais que globalement les séquences ont bien été corrigées. Le BluM montre que les corrections n'ont pas introduit de flou de manière significative dans ces séquences.

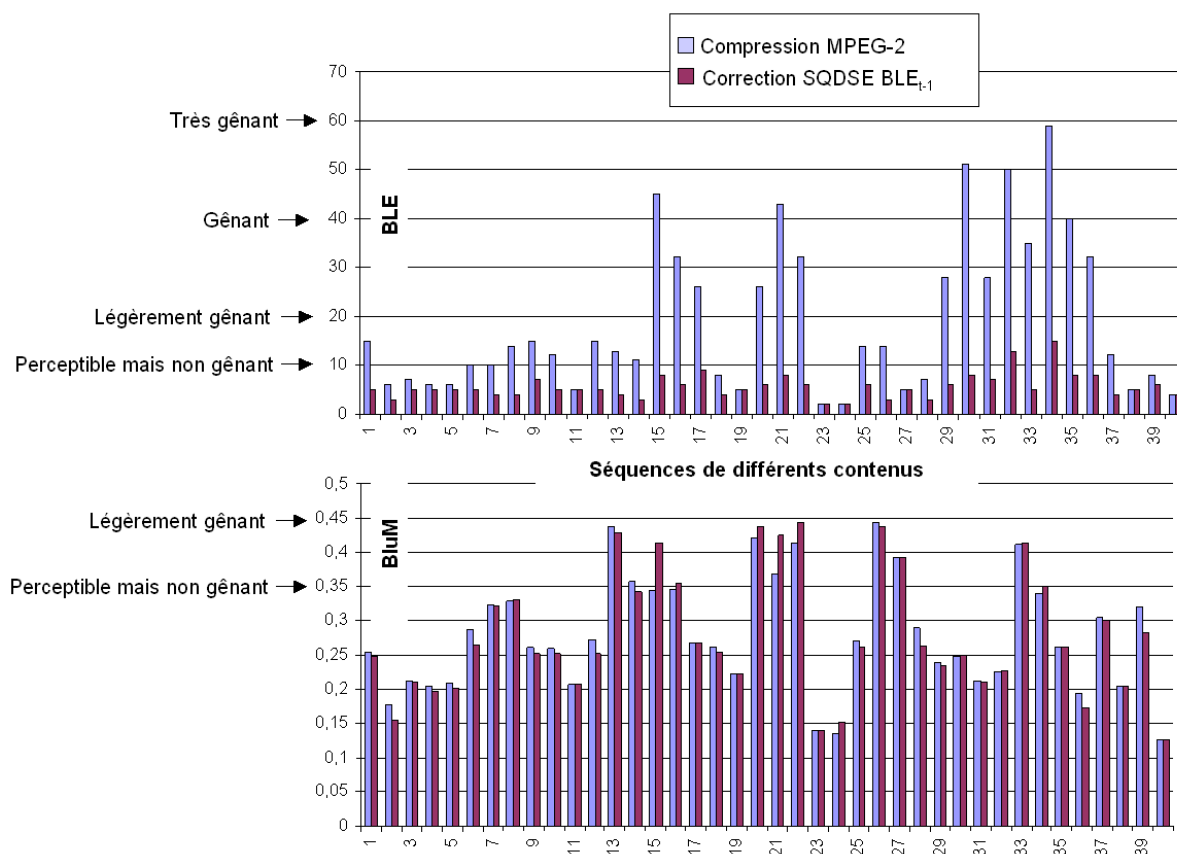


FIG. 7.18 – BLE et BluM moyen sur 40 séquences de différents contenus et plus ou moins compressées en MPEG-2 et corrigées avec le SQDSE.

L'image de la figure 7.19 présente plusieurs difficultés et voyons comment le SQDSE les contourne :

- L'image de gauche de la figure 7.20 montre des frontières de bloc fortement compressées proche d'informations capitales de l'image. La correction du SQDSE montre l'élimination des frontières de blocs et le respect du contenu de l'image.
- L'image de gauche de la figure 7.21 montre des frontières de bloc dans des zones texturées. Le SQDSE supprime ces frontières tout en préservant la texture.
- L'image de gauche de la figure 7.22 montre des frontières de bloc et du ringing proches d'un contour. Le SQDSE a supprimé l'effet de bloc et l'effet de ringing et n'a pas dégradé ou supprimé le contour. Il a reconstruit le maximum d'information par rapport aux informations disponibles de l'image compressée.

Le SQDSE offre donc la possibilité de supprimer l'effet de bloc et l'effet de ringing en préservant l'information globale de l'image, les contours et la texture. De plus le BluM indique que sans SQDSE $BluM = 0.33$ et qu'après SQDSE le $BluM = 0.35$. Cela montre que le niveau de flou n'a quasiment pas changé. En effet, la suppression des artefacts qui donne normalement une impression de flou est compensée par le rehaussement de contraste donnée par la suppression de l'effet de ringing proche des contours.



FIG. 7.19 – Image de test pour l'évaluation du SQDSE.

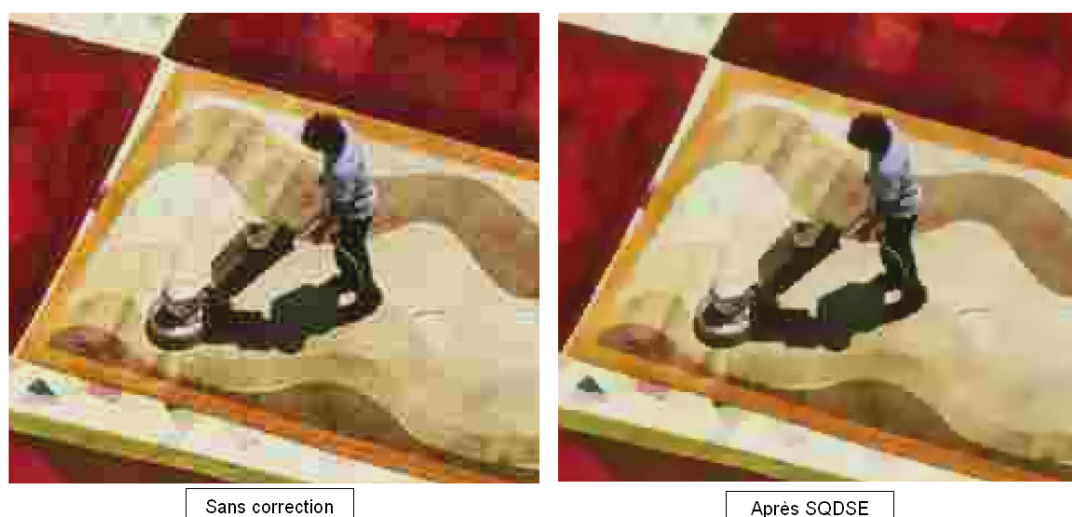


FIG. 7.20 – Efficacité du SQDSE face au respect de contenu et à la suppression de frontières prononcées.

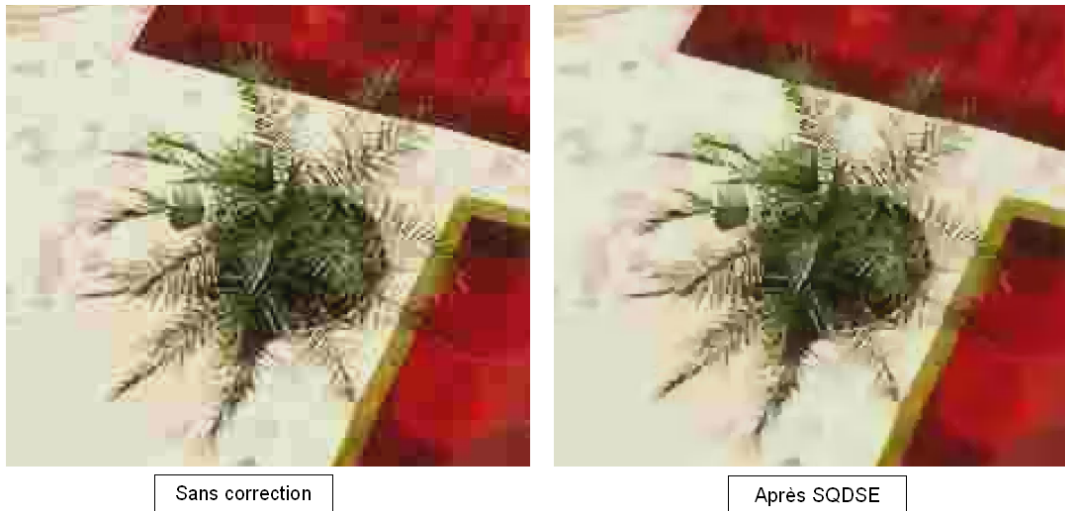


FIG. 7.21 – Efficacité du SVDSE face au respect de la texture.

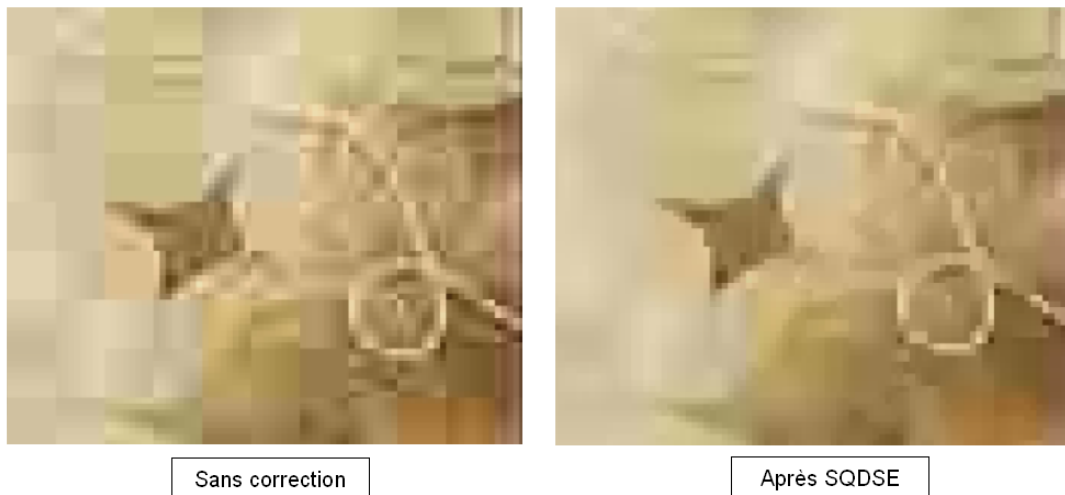


FIG. 7.22 – Efficacité du SVDSE face au respect des contours.

Nous avons jusqu'ici montré des morceaux d'images pour mettre l'accent sur l'efficacité du SVDSE sur des zones spécifiques. Les figures 7.23 et 7.24 montrent les résultats obtenus sur des images entières tirées de séquences animées : nous observons que l'effet de bloc et l'effet de ringing sont bien supprimés ou atténués quelque soit le mode codage du MBs et que ce soit pour les frontières verticales ou horizontales.

Nous venons de montrer l'efficacité du SVDSE à corriger l'effet de bloc et l'effet de ringing sans faire apparaître du flou et sans dégrader le contenu des images.



FIG. 7.23 – Image *Lilo* compressée et corrigée par le SQDSE.



FIG. 7.24 – Image *Canada* compressée et corrigée par le SQDSE.

3 Les extensions

Le SQDSE a été développé pour réduire les artefacts de compression de type MPEG-2. Nous avons vu qu'il atténuait l'effet de bloc, l'effet de ringing et aussi l'effet de mosquito. Néanmoins, pour les fortes compressions, nous allons voir quel est l'impact de la correction d'un artefact sur la perception d'un autre et plus particulièrement pour les fausses couleurs.

3.1 L'intégration d'une correction des fausses couleurs

Nous avons décidé d'axer nos recherches sur la correction d'artefacts tels que l'effet de bloc et l'effet de ringing qui nous paraissent les plus prioritaires et les plus simples à corriger au vu des contraintes imposées. Or, si on corrige l'effet de bloc et l'effet de ringing, un artefact tel que les fausses couleurs devient plus gênant car il n'est plus masqué par les autres artefacts. Ce phénomène est illustré sur la figure 7.25 avec une image compressée représentée avant et après SQDSE. L'effet de bloc et l'effet de ringing ont été supprimés mais l'image ne paraît pas pour autant de bonne qualité à cause des fausses couleurs.

Le SQDSE peut éventuellement atténuer les fausses couleurs proches des contours car, en supprimant le bruit sur la composante de luminance, cela peut atténuer le rendu créé par la combinaison de la luminance et de la teinte. De plus, de part sa légère correction en chrominance, le SQDSE peut également supprimer les frontières en couleurs. En revanche, le SQDSE ne supprimera pas les fausses couleurs. En effet, pour une correction des fausses couleurs, une analyse précise des composantes de chrominance est nécessaire [Catorina et al., 2007].

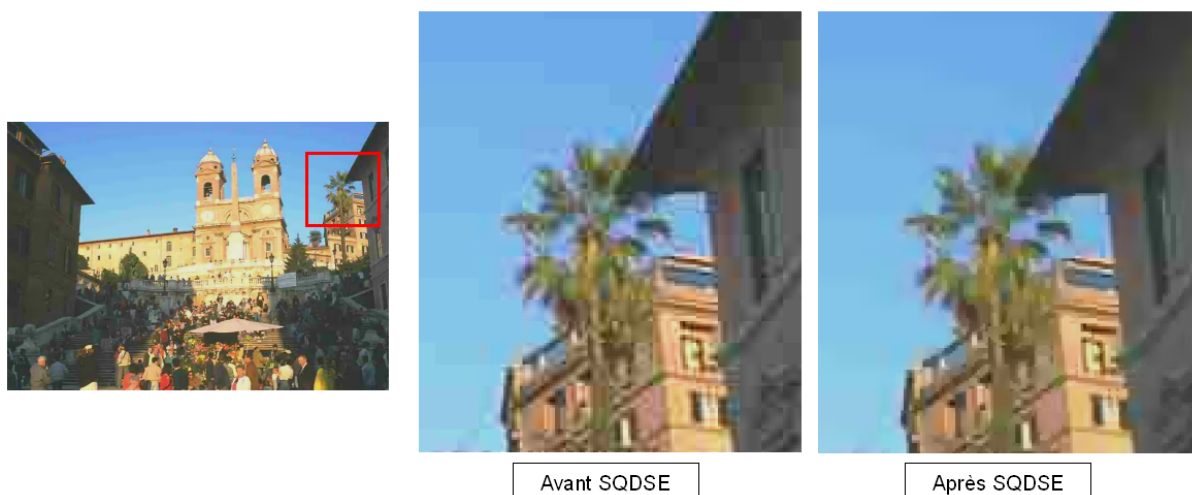


FIG. 7.25 – L'impact de la correction de l'effet de bloc et de l'effet de ringing sur la perception des fausses couleurs.

3.2 L'automatisation de la recherche de taille des MBs et du décalage

De manière à rendre le SQDSE plus autonome et pour un domaine d'application plus large, l'idéal serait d'avoir un détecteur automatique de taille de blocs. Dans le cas d'une taille identique des MBs pour toute l'image, une analyse sur l'image précédente pourrait être suffisante. En revanche, pour différencier des tailles de MBs au sein d'une même image (cas du MPEG-4/H264), l'analyse semble plus complexe si nous n'avons pas accès aux paramètres du décodeur. Dans ce cas, le SQDSE fera une correction sur des MBs de type 16 pixels par 16 pixels et ne corrigera

pas les frontières de MBs plus petits. Mais si on considère que les petites tailles de MBs sont associées aux objets en mouvement généralement encodés avec la compensation de mouvement, ces frontières ne seront pas les plus perceptibles et donc les plus importantes à corriger.

3.3 Les tests subjectifs sur des séquences

Même si nous avons vu l'efficacité du SQDSE sur chaque image d'une séquence et sa stabilité en temporel, il serait intéressant d'étudier plus précisément les phénomènes causés par cet aspect temporel. Ces phénomènes peuvent être des artefacts de compression tel que l'effet de moustique mais aussi des phénomènes sur la perception des artefacts en temporel. L'idéal serait de comparer une série de tests subjectifs sur plusieurs séquences à une série de tests subjectifs sur des images clés de ces séquences. En comparant ces deux tests, on pourrait détecter des problèmes dépendants de l'aspect temporel et étudier l'impact du temporel sur l'estimation de la qualité. Remarque : ces images clés pourraient dépendre des changements de plan et les mouvements de caméra et nous pourrions utiliser pour cela l'état de l'art sur les résumés de vidéos [Guironnet, 2006].

Pour compléter l'utilisation du SQDSE, nous avons souligné l'importance d'intégrer la correction des fausses couleurs qui deviennent beaucoup plus gênants une fois que l'effet de bloc et l'effet de ringing sont atténués. Pour le rendre plus autonome, il serait intéressant d'automatiser la recherche des tailles des MBs. Enfin, pour optimiser son utilisation, on pourrait la compléter par une étude de l'impact du temporel sur la perception des artefacts.

Dans ce chapitre, nous venons de proposer un algorithme complet de correction des artefacts de compression MPEG-2. La description de cet algorithme nous a permis d'expliquer les difficultés liées aux contraintes temps réel ou faible coût ou encore le mode de codage des MBs.

Le SQDSE est un algorithme permettant de supprimer sans référence l'effet de bloc sur les composantes de luminance et chrominance et d'atténuer l'effet de ringing et de moustique sur la composante de luminance. Son efficacité et sa robustesse sont induits par des méthodes simples guidées par une mesure de l'effet de bloc : le BLE.

La combinaison du BLE et du BluM permet de mettre en avant l'efficacité du SQDSE qui supprime les artefacts sans faire apparaître de flou. Enfin, nous avons présenté quelques extensions qui pourraient rendre le SQDSE plus autonome et plus complet.

Ce chapitre permet de répondre à la problématique du sujet en proposant une méthode capable de supprimer l'effet de bloc et l'effet de ringing sans augmenter le niveau de flou : cela est représenté sur la figure 7.26.

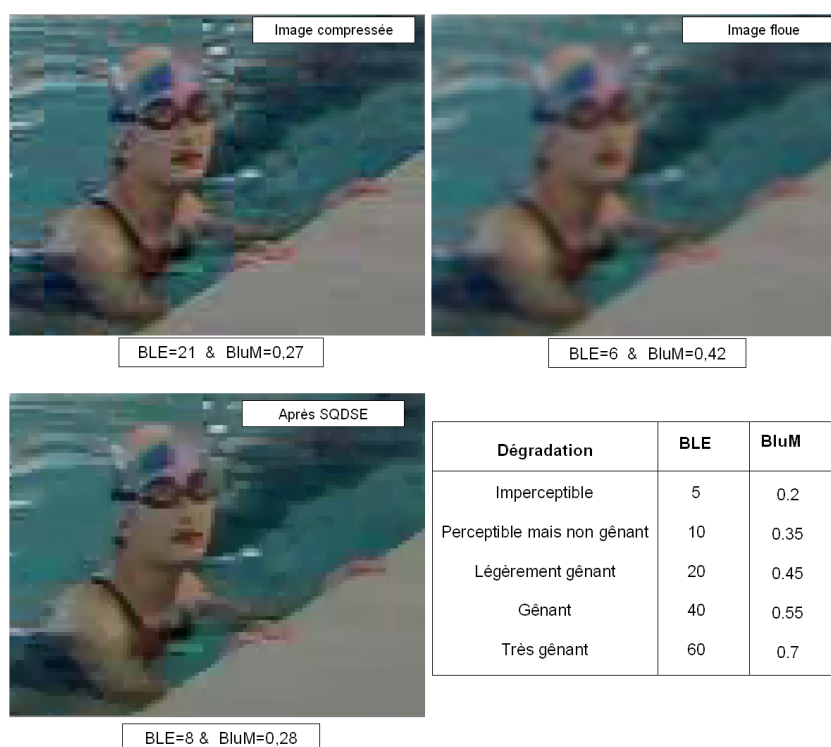


FIG. 7.26 – L'efficacité du SQDSE.

Bilan de la troisième partie

Cette troisième partie était consacrée à la correction des artefacts de compression pour la télévision numérique. Comme cela est représenté sur la figure 7.27, c'est avec l'aide de l'estimation de la qualité des différents artefacts que nous avons proposé une solution à la problématique. Le

premier chapitre de cette partie était consacré à l'explication de modules permettant d'éviter la création de fausses frontières dans le cas d'images fortement compressées, et d'éviter l'apparition du flou suite à une correction mal positionnée ou trop forte. L'efficacité de ces modules a été prouvée par des tests visuels et par nos métriques objectives validées dans la partie 2 de ce rapport. Ces modules peuvent être intégrés à d'autres algorithmes souhaitant résoudre ces problèmes.

Dans le but de proposer une solution complète, nous avons développé le SQDSE, algorithme de réduction des artefacts de compression de type MPEG-2. Après avoir expliqué comment contourner les contraintes d'implémentation, nous avons évalué son efficacité en le comparant à une autre méthode efficace et en le soumettant à nos métriques objectives. Nos métriques, ne tenant pas compte de l'aspect temporel, cette analyse a été effectuée image par image sur des séquences plus ou moins compressées.

L'étude de la formation des différents artefacts et de la gêne qu'ils occasionnent nous a donc permis dans ce chapitre, de proposer des méthodes d'amélioration de la qualité des images et des vidéos compressées, et de quantifier cette amélioration.

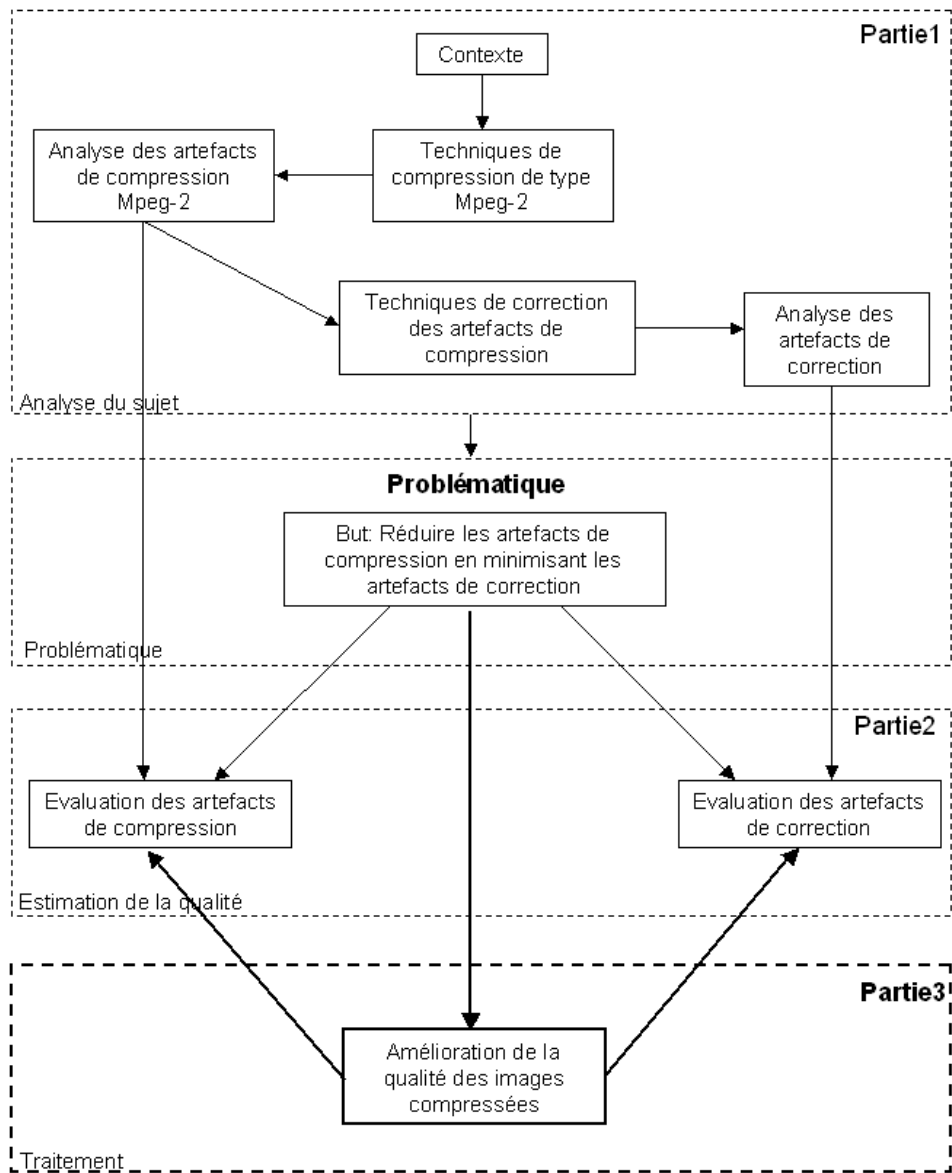


FIG. 7.27 – Bilan de la seconde partie.

Conclusion générale

A partir d'un objectif ciblé sur la réduction des artefacts de compression de type MPEG-2, nous avons abouti à l'estimation, la mesure et la correction des artefacts de compression pour la télévision numérique. Le cheminement suivi est représenté sur la figure 7.27. En suivant le travail réalisé, nous présentons nos conclusions et les perspectives qui en découlent.

Synthèse du travail et conclusions

Analyse du sujet

La compression des images est nécessaire à leur transmission, et les techniques de compression utilisées sont décrites dans des normes. L'historique de ces normes montre que la norme MPEG-2 est actuellement la plus utilisée pour la télévision numérique. L'analyse des techniques de compression utilisées dans celle-ci permet d'appréhender les artefacts causés par les pertes d'informations. Parmi ces artefacts, l'effet de bloc et l'effet de ringing sont les plus gênants et donc prioritaires à corriger. Des méthodes dédiées à leur correction sont proposées dans la littérature. Pour pouvoir les exploiter, nous sélectionnons les méthodes de classification et filtrage qui peuvent être sans référence, sans accès aux paramètres du décodeur, et implémentées à faible coût et en temps réel. La description de ces méthodes montre que ce type de correction est soit insuffisant, soit responsable de l'apparition de flou.

Cette analyse du sujet soulève la problématique suivante : comment corriger les artefacts de compression sans faire apparaître des artefacts de correction ? Pour résoudre cette problématique, nous choisissons de mesurer ces artefacts avant de les corriger.

L'évaluation de la qualité

Pour automatiser l'évaluation de la qualité des images diffusées, l'utilisation des métriques objectives sans référence semble la plus appropriée même si les tests subjectifs sont nécessaires à leur validation.

Pour évaluer les artefacts de compression, l'étude de l'effet de bloc est privilégiée. En considérant que la détection et la mesure de l'effet de bloc peuvent être utilisées pour l'étape de correction, nous ciblons les métriques de l'effet de bloc implémentables à faible coût. L'utilisation de la structure des métriques de l'état de l'art associée à quelques améliorations nous permet d'aboutir au développement de notre métrique de l'effet de bloc appelée BLE. La corrélation obtenue entre les résultats du BLE et une série de tests subjectifs valide ce BLE pour des images de contenu naturel. De plus, nous montrons les améliorations apportées par le BLE par rapport

aux autres métriques. Le BLE est donc capable de déterminer si l'effet de bloc est visible. En revanche, il n'est pas destiné à détecter les éventuelles dégradations causées par les corrections.

Pour compléter cette évaluation de la qualité, nous évaluons le flou, considéré comme le principal artefact de correction. L'étude de l'état de l'art révèle qu'il est plus judicieux de tenir compte des conséquences du flou plutôt que de ses causes et montre qu'une analyse du comportement des contours est appropriée pour détecter ce flou. En intégrant d'autres conséquences du flou sur notre perception, nous proposons une nouvelle métrique : le BluM. Des tests subjectifs dédiés à notre application et la comparaison avec les autres métriques montrent la corrélation importante du BluM avec notre perception.

Par ailleurs, une étude complémentaire sur le rôle de la couleur sur notre perception du flou montre que le développement du BluM sur la seule composante de luminance est suffisante.

A ce stade du travail, nous pouvons évaluer le niveau de bloc et le niveau de flou sur une image diffusée. Pour résoudre la problématique, il nous reste à proposer des méthodes de correction capables de délivrer une image en minimisant à la fois l'effet de bloc et l'effet de flou.

La correction des artefacts

En reprenant les limitations des corrections de l'effet de bloc et de l'effet de ringing, nous développons des modules d'amélioration pour chacune d'entre elles. L'utilisation du BLE pour contrôler la force de correction et éviter la confusion entre une frontière de bloc et un contour permet de rendre plus robuste la correction dans les zones homogènes. De plus, l'intégration de la correction des blocs étendus dans un algorithme implémentable à faible coût et en temps réel, permet aussi d'améliorer les corrections dans ces zones homogènes. Puis, nous nous intéressons aux corrections dans les zones contenant des contours ou de la texture. Ces zones sont à la fois propices à l'effet de bloc et à l'effet de ringing et la difficulté réside dans la suppression de ces artefacts sans dégrader le contenu de l'image. Pour cela, nous développons la méthode du FIV capable de distinguer automatiquement le bruit des contours et de la texture. De plus, en intégrant à cette méthode la correction de l'effet de bloc contrôlée par le BLE, nous pouvons assurer une correction de l'effet de bloc et de l'effet de ringing ne dégradant pas l'information quelque soit le contenu de l'image.

Nous intégrons ces modules d'amélioration dans une solution appelée SQDSE, destinée à corriger l'effet de bloc et l'effet de ringing sur les images qui seront diffusées. La structure de cet algorithme, son efficacité validée par des tests visuels, par le BLE et par le BluM, et sa robustesse induite par le contrôle du BLE, font du SQDSE une solution à notre problématique.

Perspectives

Des perspectives sont à envisager pour optimiser les étapes de validation, rendre encore plus autonomes le BLE, le BluM et le SQDSE et étendre leur domaine d'application.

Perspectives sur l'étude de l'effet de bloc

La détection des frontières fantômes Nous avons décrit dans ce rapport le problème des frontières fantômes et la capacité du SQDSE à les supprimer. En revanche, ni le BLE, ni les autres métriques étudiées ne sont capables de les détecter et d'appuyer les tests visuels par des notes de qualité. En considérant que nos traitements ont lieu à la réception des images, aucun traitement n'aura été effectué et l'utilisation du BLE dans son état est suffisante. En revanche, pour des compressions de type MPEG-4 AVC, la correction dans la boucle atténuera les frontières de bloc et si ces frontières ne sont pas totalement corrigées, le BLE ne pourra pas les détecter et les recorriger. L'adaptation du BLE pour le rendre capable de détecter des frontières fantômes, et sa validation par des tests subjectifs contenant à la fois des images compressées et des images corrigées, sont deux étapes nécessaires pour rendre optimale l'utilisation du BLE et du SQDSE dans le cadre d'une compression MPEG-4 AVC.

Automatiser la détection et la correction des frontières L'étude et la correction de l'effet de bloc ont été conçues pour traiter n'importe quel type d'images compressées ou non. En revanche, les résultats du BLE et du SQDSE ne seront optimaux que pour des images compressées en MPEG-2. La détection automatique des tailles de MBs dans une image pourrait d'une part rendre plus autonomes le BLE et le SQDSE face aux images redimensionnées ou décalées, et d'autre part rendre possible la correction des frontières de blocs appartenant à des tailles de MBs différentes de 16x16 pixels. Ceci optimiserait les résultats dans le cadre d'une compression de type MPEG-4 AVC qui va progressivement remplacer la compression MPEG-2.

L'analyse de notre perception des blocs en temporel Pour contrôler la correction de l'effet de bloc, le BLE se devait de traiter une séquence image par image. La validation a donc été effectuée par une étude subjective sur des images fixes. Il serait intéressant d'étudier la gêne occasionnée par l'effet de bloc lorsqu'ils sont visualisés sur des séquences. En effet, la comparaison de cette étude avec celle réalisée sur des images fixes permettrait de mettre en évidence la différence de comportement du SVH en spatial et en temporel.

Perspectives sur le flou

L'analyse de notre perception du flou en temporel Le BluM a été développé par complémentarité du BLE pour montrer que notre suppression de l'effet de bloc et de l'effet de ringing n'était pas responsable de l'augmentation du flou. Dans cet esprit, le BluM est également appliqué image par image sur une séquence. Comme pour le BLE, il serait intéressant d'analyser la différence de perception du flou sur une image fixe et sur une séquence.

La robustesse au bruit analogique Le BluM n'est pas robuste au bruit de type analogique. Notre étude étant dédiée à la correction des artefacts pour la télévision numérique, cette limitation n'est pas prioritaire. Mais, pour d'autres applications, les images peuvent être confrontées à

ce type de bruit. La modification de l'étude pixel à pixel ou l'estimation du bruit devront alors être intégrées au BluM pour le rendre plus robuste.

Perspectives sur l'effet de ringing

Le développement d'une métrique Pour évaluer la qualité des images compressées, nous avons restreint notre étude à l'effet de bloc en considérant que l'évolution de l'effet de ringing était liée à celle de l'effet de bloc. Or, la correction de l'effet de bloc n'implique pas la correction de l'effet de ringing. Donc, si seul l'effet de bloc a été corrigé, le SQDSE n'autorisera pas la correction de l'effet de ringing et son utilisation ne sera pas optimale. En revanche, une métrique de l'effet de ringing pourrait déclencher cette correction. Pour développer une telle métrique, on peut supposer utiliser l'étape de détection de l'effet de ringing utilisée pour le corriger.

La validation de notre correction de l'effet de ringing Nous avons comparé l'efficacité du SQDSE incluant une correction de l'effet de bloc et de l'effet de ringing au MSU qui corrige uniquement l'effet de bloc. Il faudrait comparer notre méthode à une autre correction d'effet de ringing. Néanmoins, nous avons prouvé l'efficacité de la correction de l'effet de ringing par des tests visuels et en montrant avec le BluM que la correction ne dégradait pas les contours. Une métrique propre à l'effet de ringing serait un plus pour valider ces résultats.

Perspective sur les fausses couleurs

La correction des fausses couleurs La correction d'un tel artefact nécessite l'étude parallèle des 3 composantes Y, Cb et Cr et certainement le changement d'espace pour pouvoir utiliser la notion de distance couleur. Le nombre et la complexité des calculs nécessaires à cette correction augmenterait la surface nécessaire du Silicium et donc le prix du circuit final. L'application visée étant réservée au domaine grand public, ce type de correction n'est donc pas prioritaire. Néanmoins, nous avons montré que la suppression de l'effet de bloc et de l'effet de ringing rendaient d'autant plus gênante la perception des fausses couleurs. Cette correction est donc à envisager pour un avenir proche.

Perspective sur la qualité

Vers une note de qualité générale Nous avons évoqué plusieurs fois l'importance de bien définir l'artefact ciblé par une métrique pour ne pas porter de conclusions trop hâtives sur la qualité générale. En supposant que nous ayons une métrique de l'effet de ringing et une métrique des fausses couleurs, il faudrait étudier comment les grouper avec le BLE et le BluM pour obtenir une note de qualité plus globale. L'intérêt de cette étude serait d'obtenir, pour une même technologie d'écran, une seule note capable d'évaluer la qualité des images diffusées. La difficulté sera alors d'attribuer un degré d'importance à chacune de ces métriques en tenant compte du fait que l'atténuation d'un artefact augmente la gêne des autres.

Glossaire

- ADSL : Asymmetric Digital Subscriber Line
- AVC : Audio Video Coding
- AVS : Audio Video coding Standard
- BLE : Block Level Estimator
- BluM : Blur Metric
- CEI : Commission Electrotechnique Internationale
- Chrominance : Terme anglais fréquemment utilisé dans la langue française pour désigner les composantes chromatiques
- CIF : Common Intermediate Format
- CODEC : Codeur Décodeur
- CSF : Contrast Sensibility Function
- DCT : Discrete Cosinus Transform
- DSCQS : Double Stimuli Continuous Quality Scale
- DVB : Digital Video Broadcasting
- FFT : Fast Fourier Transform
- FIV : Filtrage Inverse aux Variations
- GOP : Group of Pictures
- ISO : International Organization of Standardization
- ITU : International Telecommunication Union
- JND : Just Noticeable Difference
- JPEG : Joint Photographic Expert Group
- JVT : Joint Video Team
- Luminance : Terme anglais fréquemment utilisé dans la langue française pour désigner la composante d'intensité.
- LUT : Look Up Table
- MB : Macrobloc
- MPEG : Moving Picture Expert Group
- MOS : Mean Opinion Score
- MSE : Mean Squarred Error
- MSU : Moscow State University
- PSF : Point Spread Function
- PSNR : Peak Signal to Noise Ratio
- QCIF : Quarter Common Intermediate Format
- RLC : Run Length Coding
- RNIS : Réseau Numérique à Intégration de Services
- RTC : Réseau Téléphonique Commuté

- SIF : Standard Input Format
- SNR : Signal Noise Ratio
- SQDSE : Studio Quality Digital Source Enhancement
- SSCQS : Single Stimulus Continuous Quality Scale
- SSIM : Structural SIMilarity
- SVC : Scalable Video Codec
- SVH : Système Visuel Humain
- TNT : Télévision Numérique Terrestre
- TPS : Télévision par Satellite
- UQI : Universal Quality Index
- VCEG : Video Coding Experts Group
- VQEG : Video Quality Experts Group
- VHS : Video Home System

Publications et Brevets

Conférences internationales :

- F. Crété, T. Dolmière, P. Ladret, M. Nicolas, *The Blur effect : perception and estimation with a new no reference perceptual blur metric*, IS&T/SPIE Electronic Imaging Symposium, Conf. Human Vision and Electronic Imaging XII, Jan. 07, SPIE.
- F. Crété, M. Nicolas, P. Ladret, *Improvement of Deblocking Corrections on Flat Areas and Proposal for a Low-Cost Implementation*, VISAPP, International Conference on Computer Vision Theory and Applications, Mar. 07.

Colloques nationaux :

- F. Crété, P. Ladret, M. Nicolas, *Perception et Estimation du flou dans les images couleurs*, Poster for EHINC 2007 (Ecole d'Hiver sur l'Image Numérique Couleur), Poitiers, France, Jan. 07.
- F. Crété, P. Ladret, M. Nicolas, *Des métriques objectives pour estimer la qualité des images compressées*, GdR ISIS Journée codage et qualité : outils et méthodologies d'évaluation, May 07.

Brevets :

- F. Crété, M. Nicolas, P. Ladret, *Method to detect blocking artifact on flat areas for MPEG post-processing*, July 06.
- F. Crété, M. Nicolas, P. Ladret, *Method to correct blocking artifact on flat areas for MPEG post-processing*, July 06.
- G. Roffet, F. Crété, *Noise reduction in color space with high level of area detection avoiding blur effect on textured area and near edges*, June 07.

Bibliographie

- [Averbuch et al., 2005] Averbuch, A., Schclar, A., et Donoho, D. (2005). Deblocking of block-transform compressed images using weighted sums of symmetrically aligned pixels. *IEEE Transactions on Image Processing*, 14(2) :200–212.
- [Babu et Perkis, 2005] Babu, R. et Perkis, A. (2005). An hvs-based no-reference perceptual quality assessment of jpeg coded images using neural networks. In *Proceedings of IEEE International Conference on Image Processing*, volume 1, pages I–433–6.
- [Balter et al., 2006] Balter, R., Gioia, P., et L.Morin (2006). Scalable and efficient video coding using 3-d modeling. *IEEE Transactions on Multimedia*, 8(6) :1147–1155.
- [Barba, 1981] Barba, D. (1981). *Traitement numérique d'images avec critère psychovisuel de qualité*. PhD thesis, université Pierre et Marie Curie, Paris VI.
- [Beermann et Wien, 2004] Beermann, M. et Wien, M. (2004). Application of the bilateral filter for quality-adaptive reconstruction. *ISO/IEC JTC1/SC29/WG11 MPEG 2004 11713*.
- [Beermann et Wien, 2005] Beermann, M. et Wien, M. (2005). Joint reduction of ringing and blocking. *ISO/IEC JTC1/SC29/WG11 M 12640*.
- [Bellaïche, 2003] Bellaïche, P. (2003). *Les secrets de l'image vidéo*.
- [Bourdon et al., 2003] Bourdon, P., Augereau, B., Olivier, C., et Chatellier, C. (2003). Une méthode de restauration d'images compressées jpeg2000 basée sur les edp. 8ème journées d'études et d'échanges COmpression et REprésentation des Signaux Audiovisuels (CORE-SA'03), p91-94, Lyon (France), 16-17 January 2003.
- [BT.500-11, 2002] BT.500-11, I.-R. R. (2002). Methodology for the subjective assessment of the quality of television pictures. In *International Telecommunication Union*.
- [Catorina et al., 2007] Catorina, A., Spampinato, G., Bruna, A., et Capra, A. (2007). Adaptive color bleeding removal for video and still dct compressed sequences. In *Proceedings of SPIE, Digital Photography II*, volume 6502.
- [Caviedes et Gurbuz, 2002] Caviedes, J. et Gurbuz, S. (2002). No-reference sharpness metric based on local edge kurtosis. In *Proceedings of IEEE International Conference on Image Processing*, volume 1, pages 53–56.
- [Caviedes et Oberti, 2003] Caviedes, J. E. et Oberti, F. (2003). No-reference quality metric for degraded and enhanced video. In Ebrahimi, T. et Sikora, T., editors, *Visual Communications and Image Processing 2003. Edited by Ebrahimi, Touradj; Sikora, Thomas. Proceedings of the SPIE, Volume 5150, pp. 621-632 (2003)*., volume 5150 of *Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference*, pages 621–632.
- [Charrier et al., 2005] Charrier, C., Larabi, C., et H.Saadane (2005). Evaluation de la qualité des images. In *Ecole d'hiver sur l'image numérique couleur*.

- [Chou et al., 1998] Chou, J., Crouse, M., et Ramchandran, K. (1998). A simple algorithm for removing blocking artifacts in block-transform coded images. In *Proceedings of International Conference on Image Processing*, volume 1, pages 377–380vol.1.
- [Coudoux et al., 1998] Coudoux, F., Gzalet, M., et Corlay, P. (1998). Reduction of blocking effect in dct-coded images based on a visual perception criterion. *Signal Processing : Image Communication*, 11 :179–186.
- [Diday, 1971] Diday, E. (1971). La méthode des nuées dynamiques. *Revue de statistique appliquée*, 19 :19–34.
- [Dolmière, 2006] Dolmière, T. (2006). Évaluation subjective de la qualité d’images. Technical report, INPG.
- [Ebrahimi et Pereira, 2002] Ebrahimi, T. et Pereira, F. (2002). *The MPEG-4 book*. Prentice Hall PTR.
- [Elder et Zucker, 1998] Elder, J. H. et Zucker, S. W. (1998). Local scale control for edge detection and blur estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(7) :699–716.
- [Eskicioglu, 2000] Eskicioglu, A. (2000). Quality measurement for monochrome compressed images in the past 25 years. In *Proceedings of the International Conference on Acoustics, Speech and Signal Processing*, pages 1907–1910.
- [Gao et al., 2002] Gao, W., Mermer, C., et Kim, Y. (2002). A de-blocking algorithm and a blockiness metric for highly compressed images. *IEEE Transactions on Circuits and Systems for Video Technology*, 12(12) :1150–1159.
- [Genista,] Genista. <http://www.secureway.ae/brochures/pqosfacts.pdf>.
- [Girod, 1993] Girod, B. (1993). *What’s Wrong with Mean-squared Error?* MIT Press.
- [Gopinath et al., 1994] Gopinath, R., Lang, M., Guo, H., et Odegard, J. (1994). Wavelet-based post-processing of low bit rate transform coded images. In *Proceedings of IEEE International Conference Image Processing*, volume 2, pages 913–917vol.2.
- [Guironnet, 2006] Guironnet, M. (2006). *Méthodes de résumé de vidéo à partir d’informations bas niveau, du mouvement de caméra ou de l’attention visuelle*. PhD thesis, Université Joseph-Fourier - Grenoble I.
- [Hong et al., 1998] Hong, S.-W., Chan, Y.-H., et Siu, W.-C. (1998). A practical real-time post-processing technique for block effect elimination. *IEEE Transactions on Circuits and Systems for Video Technology*, 8 :4–8.
- [Hsu et Chen, 1993] Hsu, Y.-F. et Chen, Y.-C. (1993). A new adaptive separable median filter for removing blocking effects. *Consumer Electronics, IEEE Transactions on*, 39(3) :510–513.
- [Hu et Haan, 2006] Hu, H. et Haan, G. D. (2006). Low cost robust blur estimation. In *Proceedings of IEEE International Conference on Image Processing*.
- [Hu et al., 1997] Hu, J., Sinaceur, N., Li, F., Tam, K.-W., et Fan, Z. (1997). Removal of blocking and ringing artifacts in transform coded images. In *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 4, pages 2565–2568vol.4.
- [ISO, 2000] ISO (2000). Viewing conditions for graphic technology and photography. Technical report, ISO 3664.

-
- [ISO/IEC, 1993] ISO/IEC (1993). Standard mpeg-1 : 11172-2, coding of moving pictures and associated audio for digital storage media at up to about 1,5 mbit/s. Technical report.
- [ISO/IEC, 1994] ISO/IEC (1994). Standard mpeg-2 :13818-2, information and technology- generic coding of moving pictures and associated audio information. Technical report.
- [ITU, 1993] ITU (1993). Video codec for audiovisual services at p x 64 kbit/s : Recommendation h.261 à l'itu-t. Technical report.
- [ITU, 1996] ITU (1996). H263 : Video coding for low bit rate communication. Technical report.
- [ITU, 1998] ITU (1998). H263+ : Video coding for low bit rate communication : Deuxième recommandation à l'itu-t. Technical report.
- [ITU, 2000] ITU (2000). H263++ : H.263 annex u, v, w and x : Compléments de la recommandation h.263 à l'itu-t. Technical report.
- [ITU, 2004] ITU (2004). Itu-r document 6q/57-e : Draft new recommendation for subjective assessment of streaming multimedia images by non-expert viewers. Technical report, EBU, Geneva.
- [Jarske et al., 1994] Jarske, T., Haavisto, P., et Defee, I. (1994). Post-filtering methods for reducing blocking effects from coded images. In *Proceedings of IEEE International Conference on Consumer Electronics*, pages 218–219.
- [Kaup, 1997] Kaup, A. (1997). Adaptive constrained least squares restoration for removal of blocking artifacts in low bit rate video coding. In *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 4, pages 2913–2916 vol.4.
- [Kleihorst et al., 1995] Kleihorst, R., Lagendijk, R., et Biemond, J. (1995). Noise reduction of image sequences using motion compensation and signal decomposition. *IEEE Transactions on Image Processing*, 4(3) :274–284.
- [Kong et al., 2004] Kong, H., Vetro, A., et Huifang, S. (2004). Edge map guided adaptive post-filter for blocking and ringing artifacts removal. In *Proceedings of the International Symposium on Circuits and Systems*, volume 3, pages III–929–32 Vol.3.
- [Kundu, 1995] Kundu, A. (1995). Enhancement of jpeg coded images by adaptive spatial filtering. In *Proceedings of International Conference on Image Processing*, volume 1, pages 187–190 vol.1.
- [Kuo et Hsieh, 1995] Kuo, C. et Hsieh, R. (1995). Adaptive postprocessor for block encoded images. *IEEE Transactions on Circuits and Systems for Video Technology*, 5(4) :298–304.
- [Kutka et al., 1996] Kutka, R., Kaup, A., et Hager, M. (1996). Quality improvement of low-data-rate compressed video signals by pre- and postprocessing. In Ohta, N., editor, *Proceedings of SPIE, Digital Compression Technologies and Systems for Video Communications*, volume 2952 of *Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference*, pages 42–49.
- [Lee et al., 1998] Lee, Y., Kim, H., et Park, H. (1998). Blocking effect reduction of jpeg images by signal adaptive filtering. volume 7, pages 229–234.
- [Lee et Park, 2001] Lee, Y. L. et Park, H. W. (2001). Loop filtering and post-filtering for low-bit-rates moving picture coding. *Signal Processing : Image Communication*, 16 :871–890.
- [Leontaris et Reibman, 2005] Leontaris, A. et Reibman, A. (2005). Comparison of blocking and blurring metrics for video compression. In *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 2, pages 585–588.

- [Lin et al., 1992] Lin, Y., Dou, J., et Zhang, E. (1992). Edge expression based on tree structure. *Pattern Recognition*, 25 :507–517.
- [List et al., 2003] List, P., Joch, A., Lainema, J., Bjntegaard, G., et Karczewicz, M. (2003). Adaptive deblocking filter. *IEEE Transactions on Circuits and Systems for Video Technology*, 13(7) :614–619.
- [Liu et Bovik, 2002] Liu, S. et Bovik, A. (2002). Efficient dct-domain blind measurement and reduction of blocking artifacts. *IEEE Transactions on Circuits and Systems for Video Technology*, 12(12) :1139–1149.
- [Luo et Ward, 2003] Luo, Y. et Ward, R. (2003). Removing the blocking artifacts of block-based dct compressed images. *IEEE Transactions on Image Processing*, 12(7) :838–842.
- [Macq, 1992] Macq, B. (1992). Weighted optimum bit allocations to orthogonal transforms for picture coding. *IEEE Journal on Selected Areas in Communications*, 10(5) :875–883.
- [Macq et Shi, 1993] Macq, B. et Shi, H. (1993). Perceptually weighted vector quantisation in the dct domain. In *Electronics Letters*, volume 29, pages 1382–1384.
- [Marichal et al., 1999] Marichal, X., Wei-Ying, M., et Zhang, H. J. (1999). Blur dertermination in the compressed domain using dct information. In *Proceedings of IEEE International Conference on Image Processing*, volume 3, pages 386–390.
- [Marziliano et al., 2002] Marziliano, P., Dufaux, F., Winkler, S., et Ebrahimi, T. (2002). A no-reference perceptual blur metric. In *Proceedings of IEEE International Conference on Image Processing*, volume 3, pages 57–60.
- [Marziliano et al., 2004] Marziliano, P., Winkler, S., Dufaux, F., et Ebrahimi, T. (2004). Perceptual blur and ringing metrics : application to jpeg2000. *Elsevier Signal Processing : Image Communication*, 19(2) :163–172.
- [Massidda et al., 2004] Massidda, F., Perra, C., et Giusto, D. (2004). A human visual system model for no-reference digital video quality estimation. In *Proceedings of SPIE, Image Quality and System Performance*, volume 5668, pages 302–313.
- [Meier et al., 1996] Meier, T., Ngan, K., et Crebbin, G. (1996). A region-based algorithm for enhancement of images degraded by blocking effects. In *Proceedings of IEEE TENCON. Digital Signal Processing Applications*, volume 1, pages 405–408vol.1.
- [Michelson, 1927] Michelson, A. (1927). *Studies in optics*. PhD thesis, university of Chicago.
- [MSU,] MSU, V. G. <http://www.compression.ru/video/deblocking/index-en.html>.
- [Nachmias et al., 1971] Nachmias, J., Sachs, M., et Robson, J. (1971). Spatial frequency channels in human vision. *Journal of the Optical Society of America*, 61 :1176–1186.
- [Nadenau, 2000] Nadenau, M. (2000). *Integration of human color vision models into high quality image compression*. PhD thesis, EPFL.
- [Nakajima et al., 1994] Nakajima, Y., Hori, H., et Kanoh, T. (1994). A pel adaptive reduction of coding artifacts for mpeg video signals. In *Proceedings of IEEE International Conference Image Processing*, volume 2, pages 928–932vol.2.
- [Nill, 1985] Nill, N. (1985). A visual model weighted cosine transform for image compression and quality assessment. *IEEE Transactions on Communications*, 33(6) :551–557.
- [Niss, 1998] Niss, B. (1998). Prediction of ac coefficients from the dc values. *ISO/IEC JTC1/SC2/WG8 N745*.

-
- [Nosratinia, 2001] Nosratinia, A. (2001). Enhancement of jpeg-compressed images by re-application of jpeg. *Journal of VLSI Signal Processing*, 27(1-2) :69–79.
- [O' Rourke et Stevenson, 1995] O' Rourke, T. et Stevenson, R. (1995). Improved image de-compression for reduced transform coding artifacts. *IEEE Transactions on Circuits and Systems for Video Technology*, 5(6) :490–499.
- [Ong et al., 2003] Ong, E., Lin, W., Lu, Z., Yang, X., Yao, S., Pan, F., Jiang, L., et Moschetti, F. (2003). A no-reference quality metric for measuring image blur. In *Proceedings of Seventh International Symposium on Signal Processing and Its Applications*, volume 1, pages 469–472.
- [Otsu, 1979] Otsu, N. (1979). A threshold selection method from gray level histograms. *IEEE Transactions on Systems, Man and Cybernetics*, SMC 9 :62–66.
- [Ozcelik et al., 1995] Ozcelik, T., Brailean, J., et Katsaggelos, A. (1995). Image and video compression algorithms based on recovery techniques using mean field annealing. *Proceedings of the IEEE*, 83(2) :304–316.
- [Ozkan et al., 1993] Ozkan, M., Sezan, M., et Tekalp, A. (1993). Adaptive motion-compensated filtering noisy image sequences. *IEEE Transactions on circuits and systems for video technology*, 3(4) :277–290.
- [Paek et al., 1998] Paek, H., Kim, R.-C., et Lee, S.-U. (1998). On the pocs-based postprocessing technique to reduce the blocking artifacts in transform coded images. *IEEE Transactions on Circuits and Systems for Video Technology*, 8(3) :358–367.
- [Pan et al., 2004] Pan, F., Lin, X., Rahardja, S., Lin, W., Ong, E., Yao, S., Lu, Z., et Yang, X. (2004). A locally-adaptive algorithm for measuring blocking artifacts in images and videos. In *Proceedings of the International Symposium on Circuits and Systems*, volume 3, pages 925–928.
- [Pang et Tan, 1997] Pang, K. K. et Tan, T. K. (1997). Optimum loop filter in hybrid coders. *IEEE Circuits And Systems For Video Technology*, 4(2) :158–167.
- [Peli, 1990] Peli, E. (1990). Contrast in complex images. *Journal of the Optical Society of America*, 7 :2032–2040.
- [Pentland et al., 1989] Pentland, A., Darrell, T., Turk, M., et Huang, W. (1989). A simple real time range camera. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*,, pages 256–261.
- [Perra et al., 2005] Perra, C., Massidda, F., et D.D.Giusto (2005). Image blockiness evaluation based on sobel operator. In *Proceedings of IEEE International Conference on Image Processing*, volume 153, pages 389–392.
- [Pun, 1981] Pun, T. (1981). Entropic thresholding : a new approach. *Computer, Graphics and Image Processing*, 16 :210–239.
- [Quintard et Larabi, 2007] Quintard, L. et Larabi, C. (2007). Vers une métrique de qualité basée sur des critères perceptuels. In *GdR ISIS*.
- [Ramamurthi et Gersho, 1986] Ramamurthi, B. et Gersho, A. (1986). Nonlinear space-variant postprocessing of block coded images. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 34(5) :1258–1268.

- [Reeve et Lim, 1984] Reeve, H. et Lim, J. (1984). Reduction of blocking effects in image coding. *Optical Engineering*, 23 :34–37.
- [Richardson, 2003] Richardson, I. E. (2003). *H.264 and MPEG-4 Video Compression : Video Coding for Next Generation Multimedia*. Wiley.
- [Samadani et al., 2004] Samadani, R., Sundararajan, A., et Said, A. (2004). Deringing and de-blocking dct compression artifacts with efficient shifted transforms. In *Proceedings of International Conference on Image Processing*, volume 3, pages 1799–1802 Vol.3.
- [Shen et Kuo, 1997] Shen, M.-Y. et Kuo, C. (1997). Review of image postprocessing techniques for compression artifact removal. In Tescher, A. G., editor, *Proceedings of SPIE, Applications of Digital Image Processing XX*, volume 3164 of *Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference*, pages 372–382.
- [Shi et al., 2006] Shi, M., Yi, Q., et Gong, J. (2006). Blocking effect reduction based on human visual system for highly compressed images. In *Canadian Conference on Electrical and Computer Engineering*.
- [Singh et al., 2007] Singh, S., Kumar, V., et Verma, H. K. (2007). Reduction of blocking artifacts in jpeg compressed images. *Digital Signal Processing*, 17(1) :225–243.
- [Sonic,] Sonic, V. <http://www.viewsoniceurope.com/fr/products/crtgraphics/g220f.htm>.
- [Su et Mersereau, 1995] Su, J. K. et Mersereau, R. M. (1995). Post processing for artifact reduction in jpeg-compressed images. In *Proceedings of IEEE International Conference on Acoustics, Speech, Signal Processing*, volume 4, pages 2363–2366.
- [Sun et al., 2006] Sun, X., Wu, F., Li, S., et Gao, W. (2006). In-loop deblocking filter for block based video coding. In *Proceedings of International Conference on Signal Processing*, volume 1, pages 33–36.
- [Symes, 2003] Symes, P. (2003). *Digital Video Compression*. McGraw-Hill.
- [Tam et al., 1995] Tam, W., Stelmach, L., Wang, L., Lauzon, D., et Gray, P. (1995). Visual masking at video scene cuts. In *Proceedings of SPIE Human Vision, Visual Processing and Digital Display*, volume 2411, pages 111–119, San Jose.
- [Tan et Ghanbari, 2000] Tan, K. et Ghanbari, M. (2000). A multi-metric objective picture-quality measurement model for mpeg video. *IEEE Transactions on Circuits and Systems for Video Technology*, 10(7) :1208–1213.
- [Trémeau et al., 2004] Trémeau, A., Fernandez-Maloigne, C., et Bonton, P. (2004). *Image Numérique Couleur : de l'acquisition au traitement*. Dunod.
- [Van Nes et Bouman, 1967] Van Nes, F. L. et Bouman, M. A. (1967). Spatial modulation transfer in the human eye. *J. Opt. Soc. Am.*, 57 :401406.
- [VQEG, 2002] VQEG (2002). Final report from the video quality experts group on the validation of objective models of video quality assessment. Technical report, VQEG. available from <ftp.its.bldrdoc.gov>.
- [Wang, 2002] Wang, Z. (2002). Universal quality index. *IEEE Signal Processing Letters*, 9 :81–84.
- [Wang et Bovik, 2006] Wang, Z. et Bovik, A. (2006). *Modern Image Quality Assessment*. Morgan & Claypool Publishers.

-
- [Wang et al., 2000] Wang, Z., Bovik, A., et B.L, E. (2000). Blind measurement of blocking artifacts in images. In *Proceedings of IEEE International Conference on Image Processing*, volume III, pages 981–984.
- [Wang et al., 2004] Wang, Z., Bovik, A., Sheikh, H., et Simoncelli, E. P. (2004). Image quality assessment : From error measurement to structural similarity. In *Proceedings of IEEE International Conference on Image Processing*, volume 13.
- [Wang et al., 2002] Wang, Z., Sheikh, H., et Bovik, A. (2002). No-reference perceptual quality assessment of jpeg compressed images. In *Proceedings International Conference on Image Processing*, volume 1, pages I–477–I–480vol.1.
- [Watkinson, 2004] Watkinson, J. (2004). *the MPEG Handbook*.
- [Winkler, 1998] Winkler, S. (1998). A perceptual distortion metric for digital color video. In *Human Vision and Electronic Imaging, Proc. SPIE*, volume 3644, pages 175–184.
- [Winkler et al., 2001] Winkler, S., Sharma, A., et McNally, D. (2001). Perceptual video quality and blockiness metrics for multimedia streaming applications. In *Proceedings 4th International Symposium on Wireless Personal Multimedia Communications*, pages 553–556.
- [Wu et Gersho, 1992] Wu, S.-W. et Gersho, A. (1992). Improved decoder for transform coding with application to the jpeg baseline system. *IEEE Transactions on Communications*, 40(2) :251–254.
- [Xiong et al., 1997] Xiong, Z., Orchard, M., et Zhang, Y.-Q. (1997). A deblocking algorithm for jpeg compressed images using overcomplete wavelet representations. *IEEE Transactions on Circuits and Systems for Video Technology*, 7(2) :433–437.
- [Yang et al., 2001] Yang, S., Kittitornkun, S., Ghu, Y. H., Nguyen, T. Q., et Tull, D. L. (2001). Low bit rate video sequence coding artifact removal. *IEEE, workshop on Multimedia Signal Processing*, 4 :53–58.
- [Yang et al., 1993] Yang, Y., Galatsanos, N., et Katsaggelos, A. (1993). Regularized reconstruction to reduce blocking artifacts of block discrete cosine transform compressed images. *IEEE Transactions on Circuits and Systems for Video Technology*, 3(6) :421–432.
- [Yuen et Wu, 1998] Yuen, M. et Wu, H. R. (1998). A survey of hybrid mc/dpcm/dct video coding distortions. *Signal Processing*, 70(3) :247–278.
- [Zakhor, 1992] Zakhor, A. (1992). Iterative procedures for reduction of blocking effects in transform image coding. *IEEE Transactions on Circuits and Systems for Video Technology*, 2(1) :91–95.

Résumé

Résumé La compression, nécessaire à la transmission des images numériques diffusées sur les télévisions, est responsable de dégradations appelées artefacts. Pour obtenir des images de meilleure qualité, la correction des artefacts doit être efficace, robuste au large panel d'images diffusées, et simple pour que ces images soient traitées en temps réel avant leur diffusion. Pour obtenir ce compromis, notre étude se divise en 3 axes : l'estimation, la mesure et la correction. En estimant quels sont les artefacts les plus gênants et les risques encourus à les corriger, nous définissons la problématique de ce travail : comment corriger les effets de bloc et de ringing qui sont des artefacts de compression sans risquer de faire apparaître un artefact de correction tel que le flou ?

Les métriques objectives, validées par des tests subjectifs, permettent de mesurer ces artefacts. L'étude de l'état de l'art, associée à la modélisation de notre manière de percevoir ces artefacts, nous permet d'aboutir au développement du BLE (Block Level Estimator) et du BluM (Blur Metric), métriques objectives destinées à estimer respectivement l'effet de bloc et le flou sur une image.

La correction de l'effet de bloc et de l'effet de ringing peut ensuite être guidée et validée par ces métriques. Nous proposons des modules permettant d'améliorer, de rendre plus robustes et d'implémenter facilement les corrections existantes. Puis, pour résoudre la problématique, nous intégrons ces modules dans une solution générale appelée SQDSE (Studio Quality Digital Source Enhancement) destinée à supprimer les effets de bloc et de ringing. La structure du SQDSE et sa validation par des tests visuels et par les résultats du BLE et du BluM montrent que le compromis entre l'efficacité, la robustesse et la simplicité est atteint.

Mots-clés : Compression, Artefacts, Effet de bloc, Effet de ringing, Flou, Qualité, Métriques objectives, Tests subjectifs, Contrainte de faible coût, Contrainte de temps-réel, Contrainte de robustesse.

Abstract The compression of image data is mandatory for the TV broadcast but generates a number of artefacts. In order to increase the overall image-quality, each correction for a given artefact has to be efficient, reliable and should use low cost operators in order to limit the processing time prior to the image display. To reach this compromise, this study is divided in 3 parts : artefacts estimation, measurement and correction.

We define the problematic of this work by estimating the most annoying artefacts and the risks of their corrections : how can we correct the blocking and the ringing effects, that are typical compression artefacts, still limiting the blur effect which is the main artefact of the correction ? These artefacts can be measured by objective metrics which are themselves validated by subjective tests. Based on the study of state-of-the-art solutions and on our analysis of the human perception of these artefacts, we develop two new metrics, the BLE (Block Level Estimator) and the BluM (Blur Metric) which are respectively able to evaluate the blocking and the blur effects in an image.

These metrics can be used to validate and to help controlling the corrections. Consequently, we propose new modules to improve existing corrections, to increase their reliability and that can be easily plugged into current solutions. Then, to solve our initial problem, we integrate these modules within a new complete solution, named SQDSE (Studio Quality Digital Source Enhancement), for the correction of the blocking and ringing effects. Its structure and its successful validation both by visual tests and by the BluM and BLE objective quality metrics show that the SQDSE meets all our initial requirements in terms of efficiency, reliability and simplicity.

Keywords : Compression, Artefacts, Blocking effect, Ringing effect, Blur, Objective quality metric, Subjective test.