



République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

UNIVERSITE DES FRERES MENTOURI CONSTANTINE

Faculté des Sciences de la Technologie

Département d'Electronique

Laboratoire d'Automatique et de Robotique

En collaboration avec le Laboratoire L2TI de l'université Paris 13

THÈSE

Présentée pour l'obtention du diplôme de doctorat 3ème cycle LMD

Option : **Automatique et traitement du signal**

par :

DAKKAR BORHEN EDDINE

Évaluation de la qualité perceptuelle des signaux multimédias :

Évaluation multi-critère basée sur la fusion des métriques

Soutenue publiquement le : 21/12/2017

Devant le jury composé de :

Pr. F.SOLTANI	Université des Frères Mentouri Constantine	Président
Pr. F.HACHOUF	Université des Frères Mentouri Constantine	Rapporteur
Pr. K.BELARBI	École nationale polytechnique de Constantine	Examinateur
Pr. S. MESHOUL	Université Abdelhamid MEHRI Constantine	Examinatrice
Dr. S.BENIERBAH	Université des Frères Mentouri Constantine	Examinateur
Dr. Z. AHMED SEGHIR	Universitaire Abbès Laghrour Khenchela	Invité

Année Universitaire : 2017-2018

الحمد لله رب العالمين خلق اللوح والقلم، وخلق الخلق من عدم، ودبر الأرزاق
والآجال بالمقادير، وحكم وجمل الليل بالنجوم في الظلم

اللَّهُمَّ صَلِّ وَسَلِّمْ وَبَارِكْ عَلَى خَاتِمِ النَّبِيِّينَ، وَسَيِّدِ الْمُرْسَلِينَ، وَ إِمَامِ الْمُتَّقِينَ،
المبعوث رحمةً للعالمين، شفيع الأمة، وكاشف الغمة، ومُجَلِّي الظُّلْمَةِ

REMERCIEMENTS

J'adresse mes plus vifs remerciements à Madame HACHOUF Fella, professeur à l'université de Constantine, en tant que directrice de thèse qui m'a toujours guidé dans mon travail. Sa grande connaissance du domaine, ainsi que son expérience, ont joué un rôle important pour l'avancement et l'aboutissement de cette thèse. Veuillez trouver ici l'expression de ma respectueuse considération et ma profonde admiration pour toutes vos qualités scientifiques et humaines.

Je remercie vivement Prof. Faouzi SOLTANI pour avoir accepté de présider ce jury et juger ce travail.

Les professeurs Sihem MESHOUL, Khaled BELARBI et Said BENIERBAH ont accepté de juger cette thèse ; qu'ils trouvent ici l'expression de ma profonde reconnaissance.

Mes remerciements vont à messieurs Ahmed Seghir ZIANOU, Azedine BAGHDADI et Aladine CHETOUANI pour l'intérêt qu'ils ont manifesté pour ce travail et leurs critiques.

Cette thèse a été réalisée au sein de l'équipe "Image et vision" du laboratoire Automatique et Robotique en collaboration avec le laboratoire de transport et de traitement de l'information L2TI de Paris 13. Je remercie tous mes camarades de l'équipe et notamment Lamia MEZAI et Amel BOULEMNADJEL pour leurs aides.

Enfin, je remercie toutes les personnes que j'ai peut-être oubliées et qui m'ont aidé pour l'accomplissement de cette thèse.

DÉDICACE

Aucune dédicace ne saurait être assez éloquente pour exprimer ce que vous méritez pour tous les sacrifices que vous n'avez cessé de me donner depuis ma naissance, durant mon enfance et même à l'âge adulte.

Je vous dédie, maman papa, ce travail en témoignage de mon profond amour. Puisse Dieu, le tout puissant, vous préserver et vous accorder santé, longue vie et bonheur.

Je dédie ce travail aussi à mes frères, à ma sœur, à mes amis et à toute ma famille.

RÉSUMÉ

La quantification de la qualité de l'image sans référence est encore un problème difficile, en particulier lorsque différentes distorsions affectent l'image observée. Une seule métrique sans référence n'est pas en mesure d'évaluer les différentes distorsions présentées dans les images. Dans ce travail, nous nous sommes concentrés sur l'évaluation basée sur la fusion de mesures. Les métriques proposées sont construites à partir de deux étapes. Tout d'abord, un ensemble de mesures de qualité pertinentes sont utilisées pour évaluer la qualité perceptuelle de l'image. Ensuite, elles sont fusionnées en utilisant les machines à vecteur de support SVR pour dériver l'indice global de qualité.

Pour la métrique 2D, nous avons utilisé l'optimisation des essaims de particules pour sélectionner les mesures appropriées. Tandis que pour la métrique 3D, nous avons incorporé la vision binoculaire dans le processus d'évaluation.

Pour évaluer la performance des métriques proposées, différentes bases d'images ont été utilisées. Les résultats obtenus démontrent clairement que les approches proposées dépassent les méthodes de l'état-de-l'art. En outre, les approches proposées sont souples et peuvent être étendues à d'autres distorsions.

Mots clés : Qualité d'image sans référence, fusion de mesures, évaluation de la qualité perceptuelle, machine à vecteur de support, optimisation des essaims de particules, vision binoculaire.

ملخص

لا يزال القياس الكمي لنوعية الصورة دون مرجعية مشكلة صعبة، خاصة في ظل التشوهات المختلفة التي تؤثر على الصورة المشاهدة. اعتماد مقياس واحد لا يكفي لتقييم الجودة خاصة في وجود عدة أنواع من التشوهات. في هذا العمل، ركزنا في تقييمنا للجودة على مبدأ دمج مقاييس مختلفة للحصول على طريقة جديدة. حيث يتم بناء المقياس على مرحلتين:

أولاً، يتم استخدام مجموعة من المقاييس لمعرفة أو لتقييم الجودة الإدراكية للصورة.

ثانياً، يتم دمجها باستخدام نظريات التراجع من أجل الحصول على المؤشر العام للجودة.

من أجل المقاييس الثنائية الأبعاد، استعملنا خوارزمية التحيين من تنقل سرب الطيور لاختيار المقاييس المناسبة.

في حين من أجل المقياس الخاص بالصور ثلاثية الأبعاد، أدرجنا نظرية محاكاة العين البشرية في عملية التقييم.

ولتقييم أداء المقاييس المقترحة، استخدمت قواعد صور مختلفة. وتظهر النتائج التي تم الحصول عليها بوضوح أن النهج المقترح يتجاوز عديد المقاييس المقترحة سابقاً. وبالإضافة إلى ذلك، تتسم الطريقة المقترحة بالمرونة حيث يمكن تمديدها لتشمل تشوهات أخرى.

الكلمات مفتاحية:

القياس الكمي لنوعية الصورة، دمج المقاييس، الجودة الإدراكية للصورة، نظريات التراجع، خوارزمية التحيين من تنقل سرب الطيور، نظرية محاكاة العين البشرية.

ABSTRACT

Quantifying image quality without reference is still a challenging problem, especially when different distortions affect the observed image. One no-reference metric is not able to assess the different distortions presented in the images. In this work, we have focused on the assessment based on a fusion scheme of multiple distortion measures. These metrics are built in two stages. First, a set of relevant image quality assessment metrics is used to evaluate the perceptual quality. Then, a support vector regression (SVR)-based fusion strategy is adopted to derive the overall index of image quality.

For the 2D metric, we have used the particle swarm optimization to select the appropriate metrics. While for the 3D metric, we have incorporated the binocular vision in the assessment process. To evaluate the performance of the proposed metrics, different image databases have been used. The obtained results demonstrate clearly that the proposed approaches outperform the state-of-the-art NR-IQA methods. Furthermore, the proposed approaches are flexible and could be extended to other distortions.

Keywords : No reference image quality, metric fusion, perceptual quality evaluation, support vector regression, particle swarm optimization, binocular vision.

TABLE DES MATIÈRES

Table des Matières	xiii
Liste des tableaux	xvii
Table des figures	xix
Liste des abréviations	xxi
Introduction générale	1
1 Introduction à la qualité d'image	5
1.1 Introduction	6
1.2 Définition de l'évaluation de la qualité d'image	7
1.3 Types de dégradations	8
1.3.1 Le bruit	8
1.3.2 Dégradation de flou	8
1.3.3 Effets de blocs	11
1.3.4 Effet d'oscillations parasites (Ringing)	12
1.3.5 Contraste	13
1.4 Évaluation subjective	14
1.4.1 Système visuel humain (SVH)	14
1.4.2 Protocoles d'évaluation subjective	16
1.4.3 Bases d'images destinées à l'évaluation de qualité	19
1.5 Évaluation objective et performance	28
1.6 Les applications de la qualité d'image	30
1.7 Conclusion	31
2 Métriques développées pour l'évaluation de la qualité d'image : État de l'art	33
2.1 Introduction	34
2.2 Métriques avec référence (MAR-EQI)	34
2.3 Métriques avec référence réduite (MRR-EQI)	35

2.4	Métriques sans référence (MSR-EQI)	36
2.4.1	Métriques basées sur les statistiques de la scène naturelle (NSS)	37
2.4.2	Métriques basées sur les réseaux de neurones	40
2.4.3	Métriques basées sur les dictionnaires (codebooks)	41
2.4.4	Limitations des métriques sans référence existantes	43
2.5	Conclusion	43
3	Fusion des scores par régression en utilisant les séparateurs à vaste marge (SVR)	45
3.1	Introduction	46
3.2	Machines à Vecteurs de Support (Support vector regression (SVR))	47
3.2.1	Machines à Vecteurs de Support (SVR) linéaire	47
3.2.2	Machines à Vecteurs de Support (SVR) non-linéaire	50
3.3	Évaluation de la qualité d'image par fusion de scores	51
3.3.1	Scores de qualité	51
3.3.2	Fusion des scores	53
3.4	Résultats et Discussions	54
3.5	Conclusion	58
4	Évaluation générique de la qualité basée sur les essais de particules et les machine à vecteur de support	59
4.1	Introduction	60
4.2	Aperçu sur les métriques d'évaluation basées sur la fusion	60
4.3	Méthode proposée	62
4.3.1	Étape de sélection des métriques	62
4.3.2	Étape de fusion	66
4.4	Résultats et Discussions	70
4.4.1	Indépendance de la base d'apprentissage	77
4.4.2	Comparaison avec les algorithmes génétiques	78
4.5	Conclusion	80
5	La théorie du contrôle de gain et l'apprentissage automatique pour la qualité 3D	81
5.1	Introduction	82
5.2	Métriques de l'évaluation des images 3D : État de l'art	82
5.3	Évaluation de la qualité des images 3D	85
5.3.1	Différents modèles de combinaison binoculaire	85
5.3.2	Métrique 3D sans référence basée sur les machines à apprentissage et la théorie du gain	86

5.4	Résultats et discussion	89
5.4.1	Bases d'image 3D	89
5.4.2	Performance de la métrique	89
5.4.3	Expérimentations pour les images Symétriques/Asymétriques	93
5.4.4	Mesure de performance : validation croisée	95
5.4.5	Apprentissage avec différents échantillons	95
5.5	Conclusion	96
	Conclusion générale	97
	Bibliographie	99

LISTE DES TABLEAUX

TABLEAU	Page
1.1 Groupes d'images dans la base Waterloo-IVC 3D phase II	26
3.1 Coefficients de corrélation linéaire PLCC	53
3.2 Moyenne de corrélation linéaire PLCC pour 1000 itérations	55
3.3 Moyenne de corrélation de Spearman SROCC pour 1000 itérations	55
3.4 Moyenne de l'erreur quadratique RMSE pour 1000 itérations	55
4.1 Résultats obtenus pour la base LIVE	71
4.2 Résultats obtenus pour la base TID2013	72
4.3 Résultats obtenus pour la base CSIQ	73
4.4 Métriques sélectionnées pour chaque dégradation	77
4.5 Les corrélations SROCC d'indépendance (Apprentissage sur LIVE et test sur TID2013 et CSIQ)	77
4.6 Comparaison la métrique proposée avec une métrique basée sur les algorithmes génétiques	79
5.1 Les résultats de corrélation PLCC (LIVE 3D Phase I)	89
5.2 Les résultats de corrélation SROCC (LIVE 3D Phase I)	90
5.3 Les résultats de corrélation PLCC (LIVE 3D Phase II)	90
5.4 Les résultats de corrélation SROCC (LIVE 3D Phase II)	91
5.5 Résultats de la métrique proposée sur la base Waterloo-IVC 3D phase I	91
5.6 Résultats de la métrique proposée sur la base Waterloo-IVC 3D phase II	91
5.7 Résultats pour les images symétrique et asymétrique	93
5.8 La performance en utilisant le modèle d'apprentissage de la base LIVE 3D phase II.	95
5.9 Moyenne de PLCC et SROCC pour différents groupes d'apprentissage-test	96

TABLE DES FIGURES

FIGURE	Page
1.1 Représentation des différentes phases de traitement d'une image	6
1.2 L'effet de dégradation , (a)Originale, (b)Dégradée	7
1.3 Flou de défocalisation.	9
1.4 Flou de mouvement.	9
1.5 Flou dû à la compression, (a)Image compressée par le JPEG2000, (b)Image compressée par le JPEG.	10
1.6 Flou résultant du traitement.	10
1.7 Flou dû à la transmission.	11
1.8 Images compressées par JPEG : (a) Effets de blocs horizontal, (b) Effets de blocs vertical.	12
1.9 Image compressée par JPEG2000.	13
1.10 Images représentant la dégradation contraste.	13
1.11 les différentes parties de traitement dans le SVH.	14
1.12 Système optique de l'œil humain	15
1.13 Méthode à simple stimulus.	17
1.14 Méthode à double stimulus (réf et dég représente l'image de référence et dégradée respectivement).	18
1.15 Méthode Comparative.	19
1.16 Images de référence de la base LIVE.	20
1.17 Images de référence de la base TID.	21
1.18 Images de référence de la base CSIQ.	22
1.19 Images de référence de la base IVC.	23
1.20 Images de référence de la base LIVE 3D phase I.	24
1.21 Images de référence de la base LIVE 3D phase II.	25
1.22 Images de référence de la base Waterloo-IVC 3D Phase I.	26
1.23 Images de référence de la base Waterloo-IVC 3D Phase II.	27
1.24 Exemple de relation entre les scores objectifs et subjectifs.	29
2.1 Le principe des métriques avec référence réduite MRR-EQI	36
2.2 Principe général des métriques basées sur les NSS.	38

2.3	Principe général des métriques basées sur les réseaux de neurones.	40
2.4	Métriques basées sur les dictionnaires (codebook).	42
3.1	Concept du marge souple (soft marging) dans le cas linéaire.	48
3.2	Diagramme de la métrique proposée	51
3.3	Graphe des scores objectifs contre les scores subjectifs	56
3.4	Plots des scores objectifs contre les scores subjectifs pour chaque dégradation, (a)JPEG2000, (b)JPEG, (c)WN, (d)Gblur, (e) FF	57
4.1	synoptique de la métrique proposée.	63
4.2	Synoptique de la métrique 6 (BLINDS2).	67
4.3	Synoptique de la métrique 8 (CPBD).	68
4.4	Courbe du PLCC par rapport au nombre d'itérations de l'algorithme PSO.	75
4.5	Tracé des scores objectifs de la métrique proposée en fonction des scores subjectifs : (a) Base LIVE, (b) Base TID2013, (c) Base CSIQ.	76
5.1	Plot des scores de la métrique proposée et les scores subjectifs : (a) LIVE 3D phase I, (b) LIVE 3D phase II, (c) Waterloo-IVC 3D phase I, (d) Waterloo-IVC 3D phase II . . .	92
5.2	Plot des scores objectifs pour les images Sym-Asym avec les DMOS correspondant : (a) Waterloo-IVC 3D phase I, (b) Waterloo-IVC 3D phase II, (c) LIVE 3D phase II	94

LISTE DES ABRÉVIATIONS

SVH :	Système visuel humain
DCT :	Transformée en cosinus discrète
LGN :	corps genouillé latéral lateral geniculate nucleus)
CCD :	dispositif à transfert de charge (charge-coupled device)
VQEG :	Le Video Quality Experts Group
ITU :	International communication union
ACR :	Absolute category rating
ACR-HR :	Absolute category rating with hidden reference
DCR :	Degradation category rating
PC :	Pair comparison method
MEQ :	Métriques d'évaluation de qualité
DMOS :	(Différence) Score d'opinion moyen
MAR-EQI :	Métriques avec référence
MRR-EQI :	Métriques avec référence réduites
MSR-EQI :	Métriques sans référence
MSE :	La moyenne de l'erreur quadratique
PSNR :	le rapport signal sur bruit
SSIM :	L'indice de similarité structurelle
DWT :	La transformée d'ondelette discrète
GDD :	distribution Gaussienne généralisée
RLE :	Codage par plage (run-length encoding)
NSS :	Les statistiques de scènes naturelles(natural scene statistics)
ML :	Machines learning
EQI :	Évaluation de la qualité de l'image

DWT :	Transformée d'ondelette discrète
SVM :	Machines à vecteurs supports
SVR :	Régression à vecteur de support, séparateurs à vaste marge
GRNN :	réseaux généraux de régression
DBN :	le réseaux de neurones d'apprentissage profond (deep belief network)
ELM :	Extreme learning machine en anglais
ZCA :	L'analyse par composante zéro
PSO :	Optimisation par essais particulaires
PLCC :	Pearson linear correlation coefficient
SROCC :	Spearman linear correlation coefficient
RMSE :	La moyenne quadratique de l'erreur
JP2K :	Distorsion de la compression JPEG2000
JPEG :	Distorsion de la compression JPEG
WN :	Bruit blanc (white noise)
Gblur :	Floue Gaussien
FF :	Bruit du à la transmission (Fast fading)
RBF :	radial basis function
EQI-3D :	Évaluation de la qualité des images 3D
JND :	Just noticeable difference
DOG :	Difference of Gaussian
K-SVD :	L'algorithme de représentation parcimonieuse
ACP :	L'analyse en composante principale
GRNN :	Les réseaux généraux de régression
CCD :	Dispositif de transfert de charge (charge-coupled device)

INTRODUCTION GÉNÉRALE

Les images ne jouent pas seulement un rôle d'accompagnement dans notre vie : elles attirent l'attention de l'œil et facilitent la compréhension, comme dit le proverbe une image vaut mille mots. L'image représente également le contenu le plus répandu et le plus utilisé, on la trouve dans les téléphones, la télévision et sur les réseaux sociaux. En effet, l'être humain a tendance à enregistrer, répertorier ou à partager les moments qu'il aime ou qui le touche. Il est donc important de choisir des images avec une meilleure qualité.

Grâce aux avancées technologiques, l'image a pris une place prépondérante dans le quotidien des sociétés contemporaines. En effet, des équipements sophistiqués ont été développés pour l'acquisition et la diffusion. La haute résolution de ces équipements se traduit par des tailles importantes des scènes capturées. Ceci nécessite des espaces de stockage importants et fait explicitement appel aux différentes méthodes de compression.

Ce développement a engendré à la fois une amélioration de la qualité, et également un gain de temps précieux. En revanche, il est aussi la source de nouvelles problématiques qui n'existaient pas auparavant. Tout d'abord, les systèmes d'acquisition génèrent des dégradations telles que le bruit et le flou. D'autre part, il y a les artéfacts dus à la compression et à la transmission.

Prenant en compte la sensibilité du système visuel humain, qui est très exigeant en terme de qualité, il est impératif d'avoir des systèmes d'évaluation capables de mesurer les effets de la dégradation sur l'image. Le premier recours s'oriente vers les mesures subjectives où on fait appel à un groupe d'observateur qui évalue la qualité des images. Ce type d'évaluation n'est pas pratique, étant donné qu'il suggère davantage en terme de ressources et du temps.

Évaluer la qualité des images sans faire appel aux observateurs est une solution envisageable. On nomme ce type d'évaluation, les métriques objectives. Elles cherchent à remplacer l'être humain tout en maintenant des capacités d'évaluation similaires. Pour cela, des algorithmes utilisant l'information contenue dans le signal image ont été mis en œuvre. La dernière décennie a connu un grand intérêt pour la qualité objective avec des dizaines de métriques qui ont été proposées.

Les mesures objectives de qualité d'image peuvent être classées en trois groupes : les mesures de qualité avec référence, les mesures de qualité avec référence réduite et les mesures de qualité sans référence. Le classement est basé sur la présence de l'image de référence. Pour les deux premiers groupes, la totalité ou une partie de l'image est présente tandis que pour le dernier groupe, l'image de référence est totalement non disponible.

Dans le cadre de cette thèse, nous nous sommes intéressés aux métriques génériques de qualité

sans référence. D'un point de vue application, ce type de métriques est le plus intéressant. Par ailleurs, il soulève des difficultés majeures pour avoir une évaluation proche du jugement humain. Après une analyse approfondie des métriques proposées dans la littérature, nous avons constaté qu'une seule métrique n'est pas capable d'évaluer la qualité, compte tenu des différentes dégradations que peut subir l'image. De ce fait, l'idée de fusion des métriques sera adoptée. Où, différentes mesures vont être fusionnées afin de couvrir plusieurs dégradations avec une seule métrique.

La thèse est composée de 5 chapitres.

Le chapitre 1 est une introduction à la qualité d'image. Cette partie représente une introduction à l'évaluation de la qualité d'image. Une définition est donnée au début. Après, les différentes dégradations seront abordées, ainsi que les artefacts qui sont introduits par chacune. Ensuite, l'évaluation subjective et objective sont présentées. À la fin du chapitre, les domaines d'application de la qualité d'image sont rapportés avec des exemples.

Le chapitre 2 est un état de l'art des métriques développées pour l'évaluation de la qualité d'image . L'ensemble des métriques existantes dans la littérature sont revues dans ce chapitre. Une suite chronologique est suivie. Dans un premier lieu, les métriques les plus connues avec références seront présentées. Ensuite, les métriques avec références réduites seront discutées. Étant donné que cette thèse s'inscrit parmi les métriques sans références ; celles-ci seront traitées de manière plus approfondie.

Le chapitre 3 est dédié à la métrique de fusion des scores par régression en utilisant les séparateurs à vaste marge. Dans ce chapitre, une nouvelle métrique est proposée. Elle est basée sur la fusion par les des machines à vecteur de support SVR. Premièrement, la théorie des SVR est détaillée. Ensuite, la métrique proposée est présentée. À la fin du chapitre, des tests sont donnés pour corroborer l'efficacité de la métrique.

Le chapitre 4 concerne l'évaluation générique de la qualité basée sur les essais de particules et les machines à vecteur de support. Cette partie vise à prolonger l'étude du chapitre précédent en complétant la métrique. Le chapitre est organisé comme suit, une revue des métriques basées sur la fusion est présentée. Ensuite, les étapes de la métrique proposée sont données ; l'étape de sélection des métriques avec les essais de particules PSO, et l'étape de fusion par SVR. Des tests exhaustifs sont rapportés afin de présenter la performance de la métrique suggérée vis-à-vis des métriques de l'état de l'art.

Le chapitre 5 présente la théorie du contrôle de gain et l'apprentissage automatique pour la qualité 3D. Ce chapitre est dédié à l'évaluation 3D. Un état de l'art des métriques existantes est présenté en premier, suivi par la métrique proposée où le modèle de combinaison

binoculaire est abordé. L'apprentissage par SVM est utilisé pour la fusion des scores objectifs. Les tests sont conduits sur plusieurs bases d'images stéréoscopiques. La métrique donne des résultats significatifs et prometteurs.

Le mémoire se termine par une conclusion et des perspectives pour d'autres éventuels axes de recherches.

INTRODUCTION À LA QUALITÉ D'IMAGE

Sommaire

1.1	Introduction	6
1.2	Définition de l'évaluation de la qualité d'image	7
1.3	Types de dégradations	8
1.3.1	Le bruit	8
1.3.2	Dégradation de flou	8
1.3.3	Effets de blocs	11
1.3.4	Effet d'oscillations parasites (Ringing)	12
1.3.5	Contraste	13
1.4	Évaluation subjective	14
1.4.1	Système visuel humain (SVH)	14
1.4.2	Protocoles d'évaluation subjective	16
1.4.3	Bases d'images destinées à l'évaluation de qualité	19
1.5	Évaluation objective et performance	28
1.6	Les applications de la qualité d'image	30
1.7	Conclusion	31

1.1 Introduction

L'image représente un moyen de communication incontournable de nos jours. Elle utilise les différents signes visuels pour pouvoir échanger l'information. Elle permet aussi de mettre en valeur des objets, des scènes,... etc.

Aujourd'hui, créer un profil sur les réseaux sociaux (Facebook, Twitter,...), partager des souvenirs avec son entourage sont des utilisations banales de l'image. Ainsi, l'utilisation du contenu visuel a connu un fort développement qui tend à s'accroître dans les prochaines années. De ce fait, le domaine de l'image a connu un élargissement considérable à la fois en termes de matériel et de techniques de traitement d'images.

Depuis l'invention des capteurs CCD dispositif à transfert de charge (charge-coupled device) en 1969 ; l'évolution des moyens d'acquisition d'images ne cessent pas d'accroître. Sur le marché, il existe une multitude de choix d'appareils qui garantissent une capacité et une meilleure qualité d'image. En parallèle, les techniques de traitement ont connu la même vitesse de progrès dans tout ce qui concerne la chaîne de traitement (acquisition, compression, transmission, diffusion,...). Au regard de cet essor majeur et la demande grandissante de l'utilisation des images, une meilleure qualité du contenu visuel est fortement exigée. Durant la chaîne de traitement Figure 1.1 et à chaque niveau, l'image subit différentes dégradations/distorsions qui peuvent nuire à sa qualité perceptuelle.

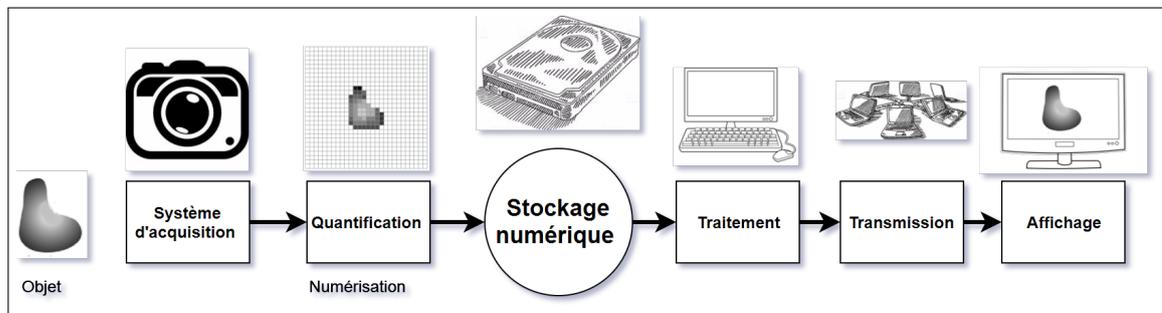


FIGURE 1.1: Représentation des différentes phases de traitement d'une image

Étant donné que l'être humain est doté d'un système visuel sophistiqué capable d'évaluer la qualité perceptuelle de n'importe quelle image et afin de satisfaire les exigences de ce système la qualité du contenu visuel doit être contrôlée impérativement avant qu'elle arrive à ce dernier. Cela peut être réalisé avec différentes manières et il est appelé évaluation de la qualité perceptuelle.

1.2 Définition de l'évaluation de la qualité d'image

La qualité d'image est une opération omniprésente dans notre vie grâce à notre système visuel humain SVH. Le fait de capturer les défauts dans une image est appelé évaluation de qualité. La Figure 1.2 représente deux images. L'une dite originale, il est supposé qu'elle ne contient pas de distorsion. La deuxième est une image déformée avec une dégradation. Notre SVH détecte facilement laquelle des deux images a la meilleure qualité.

En traitement d'image, la qualité est un domaine qui s'intéresse à la mesure de la perception des dégradations, distorsions, artefacts introduites dans une image par les systèmes de traitement. Il existe deux types d'évaluation ; subjective et objective.

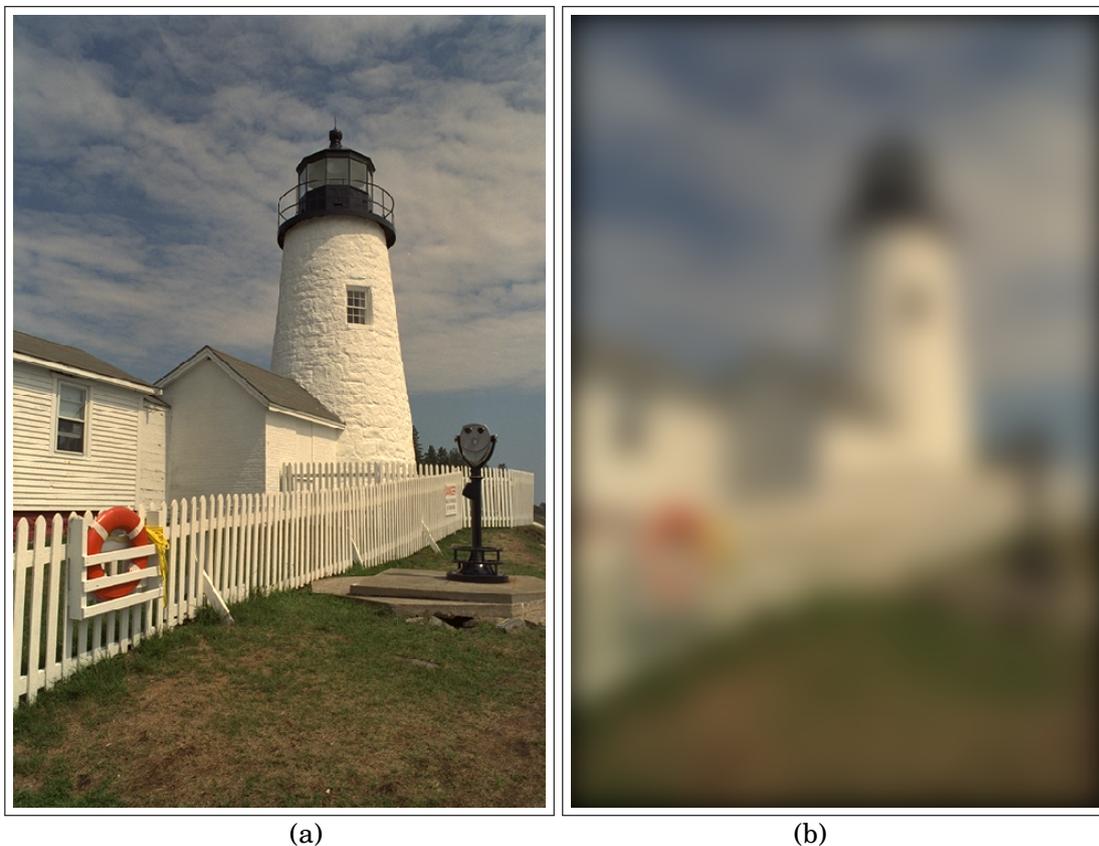


FIGURE 1.2: L'effet de dégradation , (a)Originale, (b)Dégradée

L'évaluation subjective consiste à identifier comment la qualité d'une image est perçue par les humains. Elle s'intéresse à la façon dont un observateur peut donner son avis sur une image particulière [99]. Tandis que l'évaluation objective consiste à concevoir des modèles mathématiques capables de prédire la qualité d'une image avec précision et d'une façon automatique. La métrique objective devra pouvoir imiter les prévisions de qualité données par un observateur humain [58].

1.3 Types de dégradations

On appelle distorsion ou dégradation tout défaut pertinent qui peut être vu, perçu ou remarqué dans une image. Dans la pratique, il existe plusieurs types de dégradations. Les sources distorsions se diffèrent ; il y a celles qui sont introduites durant la phase d'acquisition, d'autres sont dues aux traitement ... etc. Dans ce paragraphe, un résumé des dégradations les plus connues est donné.

1.3.1 Le bruit

Le bruit représente une dégradation courante dans l'image. Il est défini comme toutes informations parasites qui s'ajoutent à l'image, il est caractérisé par l'apparition de grains visibles. Parmi les conséquences de cette dégradation il y a la perte de netteté dans les détails de l'image. Le bruit s'ajoute à l'image selon deux formes additif ou multiplicatif [55].

- Additif

$$(1.1) \quad g(x, y) = f(x, y) + \eta(x, y)$$

- Multiplicatif

$$(1.2) \quad g(x, y) = f(x, y) \times \eta(x, y)$$

avec g l'image dégradée, f l'image originale et η le bruit.

1.3.2 Dégradation de flou

Le flou est une dégradation qui se manifeste au niveau des contours. Elle affecte les bords des objets dans l'image. Cette distorsion rend l'image inaperçue et difficilement utilisable. En qualité d'image, le flou est caractérisé par le fait qu'il dégrade les informations à haute fréquence présentes dans l'image [15], [92]. Tout au long de la chaîne, on peut avoir différents types de flou.

- **Défocalisation (Out-of-focus)** : La défocalisation est définie par la création d'une image lorsque les rayons de lumière traversant une lentille tombent sur un plan devant ou au-delà du point où elles convergent pour former une image nette. Les images hors focus apparaissent floues.

Ce type de flou se crée au moment de la capture. En effet, la focalisation sur un plan entraîne automatiquement un flou au niveau des objets qui se situent en dehors du champ de focalisation. Il est souvent modélisé par une convolution de l'image (ou une partie de l'image) avec un noyau gaussien, où sa variance est associée à la quantité de flou. La Figure 1.3 montre une défocalisation lors d'une prise de vue.



FIGURE 1.3: Flou de défocalisation.

- **Flou de mouvement** : Le flou de mouvement dit aussi cinétique désigne le flou dû au mouvement rapide des objets photographiés pendant l'enregistrement. Le flou de mouvement est formé soit par le mouvement de l'objet quand l'appareil de capture est fixe, soit par la scène à filmer si l'appareil suit l'objet en mouvement. Ce type dégrade des directions spécifiques dans le domaine fréquentiel. La Figure 1.4 présente le flou de mouvement.



FIGURE 1.4: Flou de mouvement.

- **Flou dû à la compression** : Les processus de compression représentent une source importante du flou. En général, des filtres passe-bas sont appliqués sur l'image, ils opèrent sur les hautes fréquences. Cela induit une pertes de détails et de netteté. La Figure 1.5 représente les effet des deux compressions JPEG2000 et JPEG.

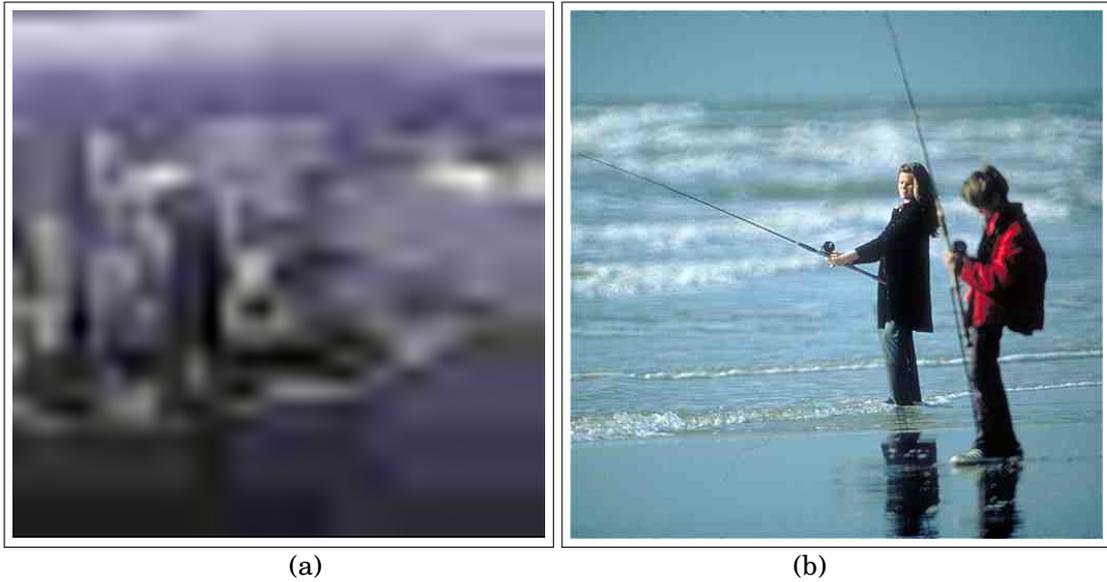


FIGURE 1.5: Flou dû à la compression, (a)Image compressée par le JPEG2000, (b)Image compressée par le JPEG.

- **Flou dû au traitement** : Durant la phase de traitement, l'image peut acquérir différentes types de dégradations. Le traitement qui est souvent responsable de la distorsion floue est le filtrage. La Figure 1.6 présente une image filtrée par un filtre Gaussien.



FIGURE 1.6: Flou résultant du traitement.

- **Flou dû à la transmission :** La transmission des images dans les canaux produit souvent des pertes d'informations, ces pertes se traduisent par des régions floues. La Figure 1.7 représente une image résultante de la simulation d'une transmission dans les canaux Rayleigh.



FIGURE 1.7: Flou dû à la transmission.

1.3.3 Effets de blocs

Dès son apparition dans les années 90, l'utilisation de la norme JPEG ne cesse pas à croître. Grâce à son efficacité de compression un gain important en bande passante peut être atteint. La majorité des appareils photos et des téléphones portables utilisent cette norme pour enregistrer les images capturées. Il faut plusieurs étapes pour coder une image en JPEG. La première étape consiste à transformer les couleurs de l'image du modèle colorimétrique initial vers un modèle de type chrominance/luminance. Un sous-échantillonnage est appliqué ensuite afin de réduire la taille de l'image. Cependant, une perte d'information irrécupérable peut être enregistrée dans cette étape. Pour chaque bloc, la transformée de cosinus discrète DCT est calculée. La DCT peut créer une perte d'informations à cause des différentes erreurs d'arrondis. A la fin de cette étape, les basses et les hautes fréquences sont distinguées. Les basses fréquences représentent les informations pertinentes dans l'image tandis que les hautes fréquences représentent les zones à fort contraste. C'est sur les hautes fréquences que la compression s'effectuera. Une fois les coefficients de DCT obtenus, c'est l'étape de quantification qui procèdera. C'est là où se produisent les majeures parties de la perte d'information et de la qualité visuelle. Le but est d'atténuer les hautes fréquences très peu sensibles par le système visuel humain SVH. La phase finale de la compression JPEG consiste à effectuer un codage en zigzag suivi par une compression selon l'algorithme de codage par plage RLE (run-length encoding) puis un codage entropique de type Huffman.

Comme la compression JPEG est basée sur un découpage par bloc de l'image, elle est responsable

de l'une des dégradations les plus nuisibles pour la qualité visuelle de l'image. On parle de l'effet de bloc ou de bords. Elle apparaît en général au niveau des frontières horizontales et verticales. La Figure 1.8 montre les effets de blocs horizontaux et verticaux.

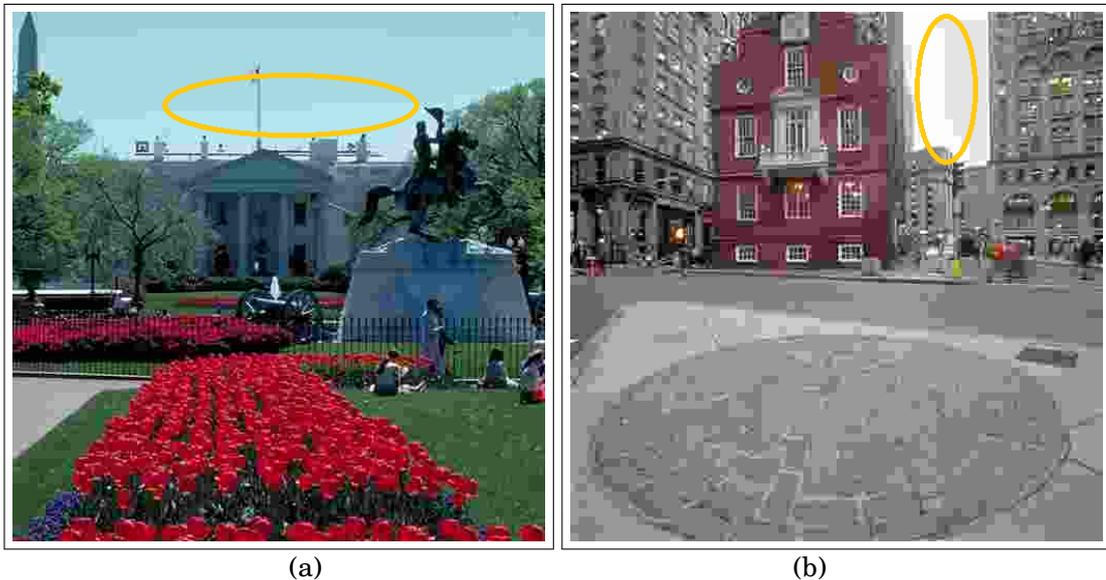


FIGURE 1.8: Images compressées par JPEG : (a) Effets de blocs horizontal, (b) Effets de blocs vertical.

1.3.4 Effet d'oscillations parasites (Ringing)

La Figure 1.9 montre deux images compressées par la compression JPEG2000. On peut clairement voir des oscillations autour des contours, cela est dû à la compression. En effet, lorsqu'une image est trop compressée, on demande à l'ordinateur de travailler avec une résolution grossière. Dans ce cas, l'ordinateur fait ce qu'il peut pour recréer les interfaces contrastées éventuellement présentes dans l'image originale. Ce qui va faire apparaître des halos vibrant (Figure 1.9) autour des contours. Ce phénomène est appelé le phénomène de Gibbs ou le ringing. En traitement d'image, le phénomène de Gibbs se manifeste après l'étape de quantification (la compression JPEG2000) sous forme d'oscillations autour des régions à forte contraste.



FIGURE 1.9: Image compressée par JPEG2000.

1.3.5 Contraste

Le contraste est un facteur déterminant dans la perception de la qualité visuelle. La distorsion du contraste est souvent introduite durant la phase d'acquisition. Les limitations du matériel d'acquisition et les conditions lumineuses sont les causes principales des pertes de contraste et de la visibilité des détails de l'image. Il existe d'autres sources de la distorsion contraste. Parmi ces sources on peut citer le processus de rehaussement qui crée une perte ou une sur-contraste [21]. La Figure 1.10 montre l'effet de contraste sur la qualité perceptuelle.



FIGURE 1.10: Images représentant la dégradation contraste.

Il existe d'autres types de dégradations dans la littérature qui sont moins connues. Le lecteur peut être orienté à la référence [67] pour plus d'informations.

1.4 Évaluation subjective

L'évaluation subjective de qualité représente toutes expérimentations psychovisuelles pour lesquelles un nombre d'observateurs évaluent un groupe de stimuli donné. A la fin, une note de qualité est attribuée à chaque image. Ces tests coutent chères en termes de temps et de ressources humaines. Mais ils restent le moyen le plus fiable d'évaluation de qualité étant donné que l'observateur humain reste toujours le récepteur final du contenu visuel.

1.4.1 Système visuel humain (SVH)

Le système visuel humain est le moyen qui assure notre vision. De point de vue traitement de l'information, il est divisé en quatre parties : traitement optique, traitement de la rétine, traitement du corps genouillé latéral LGN (lateral geniculate nucleus) et le traitement cortical. La Figure 1.11 montre les différentes parties de traitement dans le SVH.

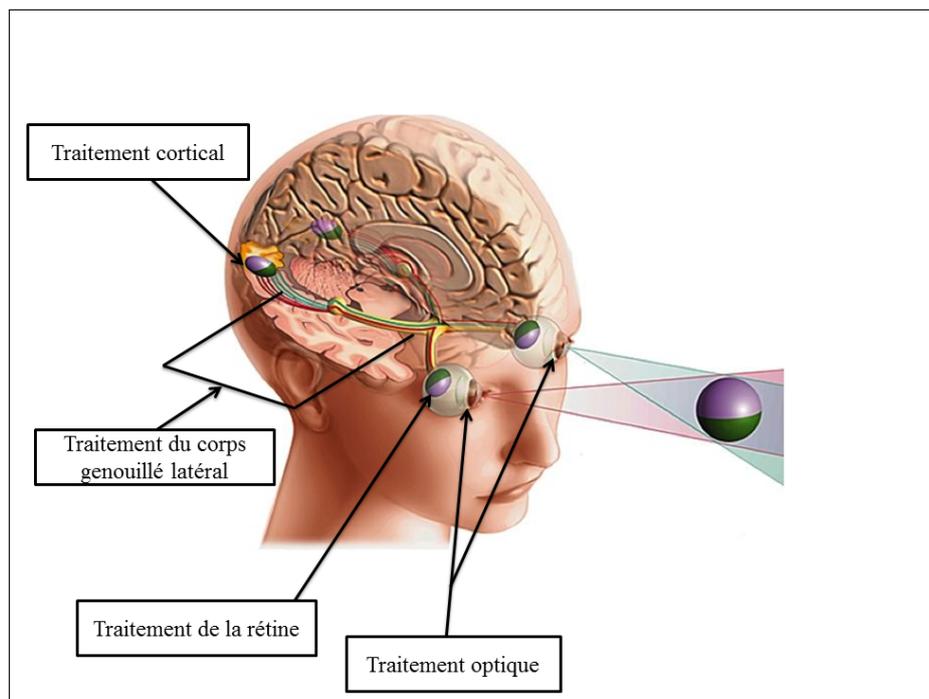


FIGURE 1.11: les différentes parties de traitement dans le SVH.

Dans la première étape, une image visuelle sous forme de lumière traverse l'optique de l'œil et est projetée sur la rétine, une membrane à l'arrière des yeux. Le résultat est nommé l'image rétinienne. L'optique est composé de trois éléments majeurs : la cornée, le pupille et l'iris [99] comme le montre la Figure 1.12. La rétine est composée de plusieurs couches de neurones. La première

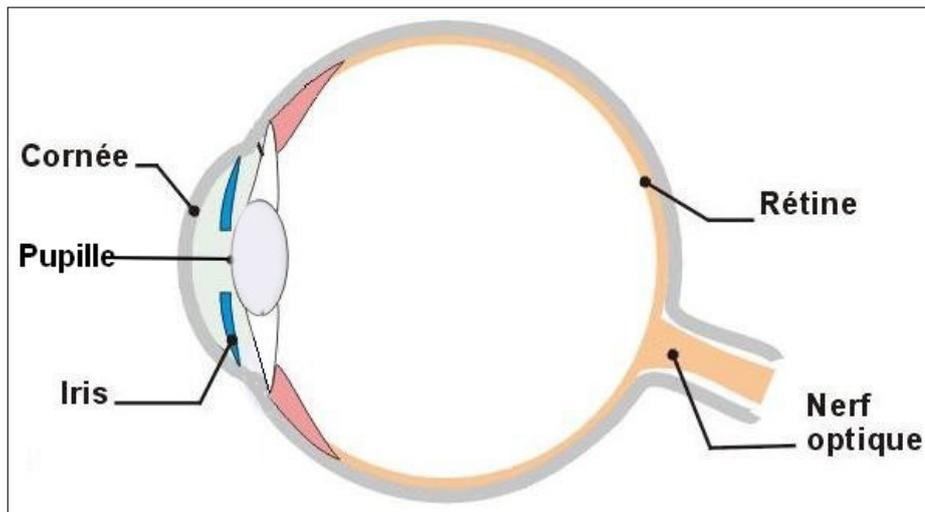


FIGURE 1.12: Système optique de l'œil humain

couche se compose de cellules photoréceptrices qui échantillonnent l'image rétinienne projetée. Il existe deux types de photorécepteurs : les cônes et les tiges (rods). Les cônes sont responsables de la vision dans des conditions de lumière élevée, tandis que les tiges sont responsables de la vision dans des conditions à faible luminosité. Il existe trois types de cônes classés selon leur spectre de sensibilité des longueurs d'ondes lumineuses : les L-cônes, les M-cônes, et le S-cônes. Ils correspondent aux ondes : longues (570 nm), moyennes (540 nm) et courtes (440 nm) pour lesquelles la sensibilité est maximale.

La distribution spectrale de la longueur d'onde (ou l'information couleur) de la lumière entrante est codée par les activités relatives des L-, M- et S-cônes. Toutes les tiges (rods) ont à peu près la même sensibilité spectrale (pics à 500 nm) et ils ne codent pas l'information couleur. Le signal discret échantillonné provenant des photorécepteurs passe par plusieurs couches de neurones interconnectés (y compris les cellules bipolaires, les cellules amacrines et cellules horizontales) avant d'être transmis aux cellules ganglionnaires, dont les fibres nerveuses (axones) forment le nerf optique : l'unité de sortie de la rétine.

Les distribution spatiale des cône photorécepteurs et des cellules ganglionnaires sont hautement non uniformes. Le point sur la rétine qui se trouve sur l'axe visuel s'appelle fovea, qui a la plus haute densité de cônes et les cellules ganglionnaires. La densité diminue rapidement en fonction de la distance du fovea, et la distribution des cellules ganglionnaires diminue plus rapidement que celle des récepteurs de cône. L'effet de ces distributions non uniformes c'est quand l'observateur

humain fixe son regard sur un point dans une image du monde réel, une image de résolution variable est transmise par le canal visuel aux unités de traitement de niveau supérieur après la rétine. En d'autres termes, lorsque vous lisez une phrase dans un texte, les phrases autour seront floues ou avec une résolution réduite. La région autour du point de fixation est projetée sur la fovea, où l'image rétinienne est échantillonnée avec la plus haute densité et perçue avec la plus haute résolution. La résolution de l'image perçue diminue rapidement avec la distance du point de fixation.

L'information codée dans la rétine est transmise via le nerf optique au LGN, avant d'être transmise au cortex visuel. Le nombre total des neurones de sorties du LGN pour le cortex visuel est légèrement plus grand que le nombre de cellules ganglionnaires connectées au LGN. Le LGN est également l'endroit où l'information de l'œil gauche et droit est fusionnée. Le cortex visuel est divisé en plusieurs couches. Parmi ces couches, le cortex visuel primaire (ou la couche V1) est directement connecté avec le LGN et contient environ 1.5×10^8 neurones, beaucoup plus que les 10^6 neurones dans le LGN. Il est constaté qu'un grand nombre de neurones dans le cortex visuel primaire sont accordés à des stimuli visuels avec une localisation spatiale spécifique, une fréquence et une orientation. Les champs réceptifs de ces neurones sont bien décrits à l'aide des passe-bandes localisées, et des fonctions orientées, qui coïncident avec le concept d'ondelettes dans la littérature de traitement du signal et les analyses harmoniques. Il existe deux types de cellules dans le cortex visuel primaire : cellules simples et cellules complexes. La classification est basée selon le type de sommation (linéaire ou pas) de la contribution à travers les neurones du champ réceptif.

1.4.2 Protocoles d'évaluation subjective

Étant donné que l'être humain est l'ultime juge de la qualité perceptuelle. Les tests subjectifs restent un moyen fiable et incontournable pour évaluer les métriques objectives. Pour construire une base d'images qui peut servir à l'évaluation de la performance des mesures proposées certaines règles doivent être respectées. Le Video Quality Experts Group (VQEG), un groupe formé par des experts de l'union internationale de communication (International communication union ITU) ont mis un certain nombre de recommandations pour mieux évaluer la qualité des images/vidéos subjectivement [35].

1.4.2.1 Note de catégorie absolue (Absolute category rating (ACR))

La méthode de notation de catégorie absolue appelée aussi la méthode à **simple stimulus** est un jugement de catégorie où les séquences de test sont présentées une par une et sont évaluées indépendamment dans une échelle de catégorie. La méthode spécifie qu'après chaque présentation, on demande aux observateurs d'évaluer la qualité de la séquence montrée. La Figure 1.13 montre le déroulement de test avec les temps nécessaires pour chaque étape [35]. L'échelle de cinq niveaux suivante est utilisée pour évaluer la qualité des stimulus :

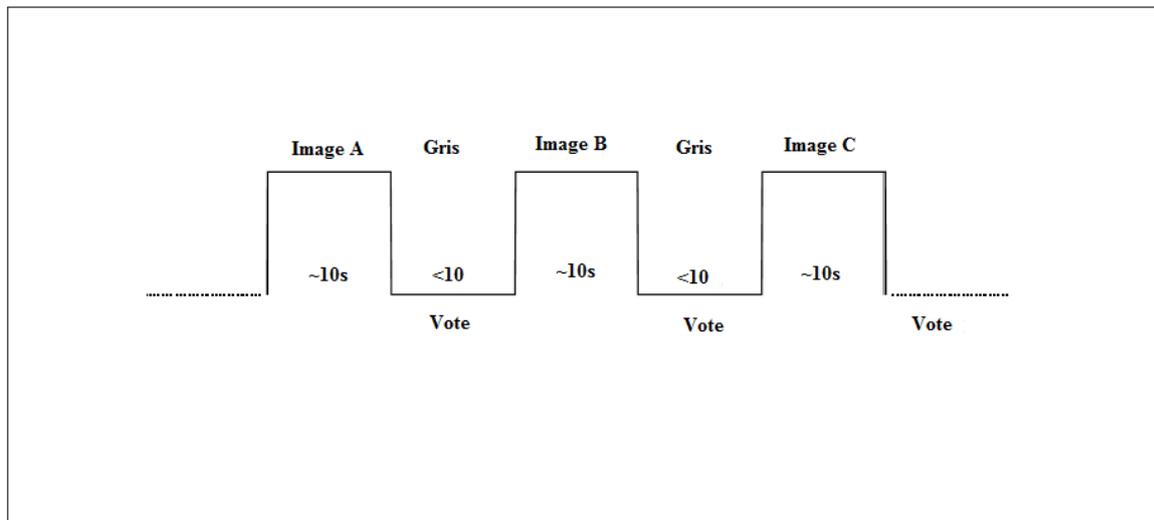


FIGURE 1.13: Méthode à simple stimulus.

- Excellente
- Bonne
- Assez bonne
- Médiocre
- Mauvaise

La note Excellente indique une meilleure qualité.

1.4.2.2 Note de catégorie absolue avec référence cachée (Absolute category rating with hidden reference (ACR-HR))

Dans cette méthode, un jugement est mis pour les séquences de test où elles sont présentées une par une et sont évaluées de manière indépendante sur une échelle. Ce test d'essai doit inclure une version de la référence de chaque séquence de test présentée comme tout autre stimulus de test. La méthode précise que, après chaque présentation, les observateurs sont invités à évaluer la qualité de la séquence montrée.

Cette méthode doit être utilisée avec une vidéo de référence où un expert du domaine l'a considéré de **bonne** ou d'**excellente** qualité. Elle est aussi incapable d'analyser les dégradations inhabituelles qui se produisent dans la première et la dernière seconde de la séquence vidéo [35].

1.4.2.3 Notation de dégradation catégorique (Degradation category rating (DCR))

Appelée aussi méthode à **double stimulus** elle implique que les séquences de test sont présentées par paires : la première séquence de chaque paire est toujours la référence, tandis que la seconde est la même mais en utilisant sa version dégradée.

La Figure 1.14 montre la méthode de présentation des différents stimulus ainsi que les temps

nécessaires.

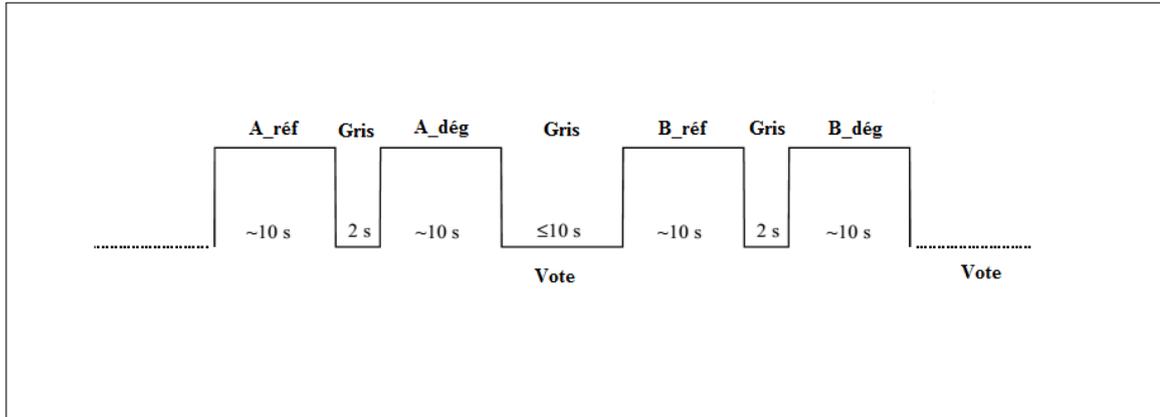


FIGURE 1.14: Méthode à double stimulus (réf et dég représente l'image de référence et dégradée respectivement).

Ici, les observateurs évaluent la qualité du second stimulus par rapport à la référence [35]. L'échelle à cinq niveaux suivante est utilisée pour évaluer les différentes distorsions :

- Imperceptible
- Perceptible
- Peu dégradée
- Dégradée
- Très dégradée

1.4.2.4 Méthode comparative (Pair comparison method (PC))

La méthode comparative fonctionne comme suit ; les séquences de test sont présentées par paires, les séquences sont présentées sur un premier système de test et ensuite sur un deuxième qui est totalement différent [35].

Les systèmes sous test (A, B, C,...,etc.) sont généralement testés en combinant les différentes possibilités AB, BA, CA,... . Après la présentation de chaque paire, un jugement est effectué sur le stimulus le plus préféré. La Figure 1.15 montre le déroulement de ce test.

avec : A_i, A_j la paire sous les i^{eme}, j^{eme} conditions respectivement, et B_k, B_l la paire sous les k^{eme}, l^{eme} conditions respectivement.

Afin d'avoir la base d'image qui représente le mieux la qualité visuelle perçue par notre SVH plusieurs facteurs doivent être pris en considération. On peut citer les conditions de visions, le choix des observateurs, les images utilisées,... etc.

Pour mieux connaître tous ces facteurs, le lecteur est orienté vers les recommandations données

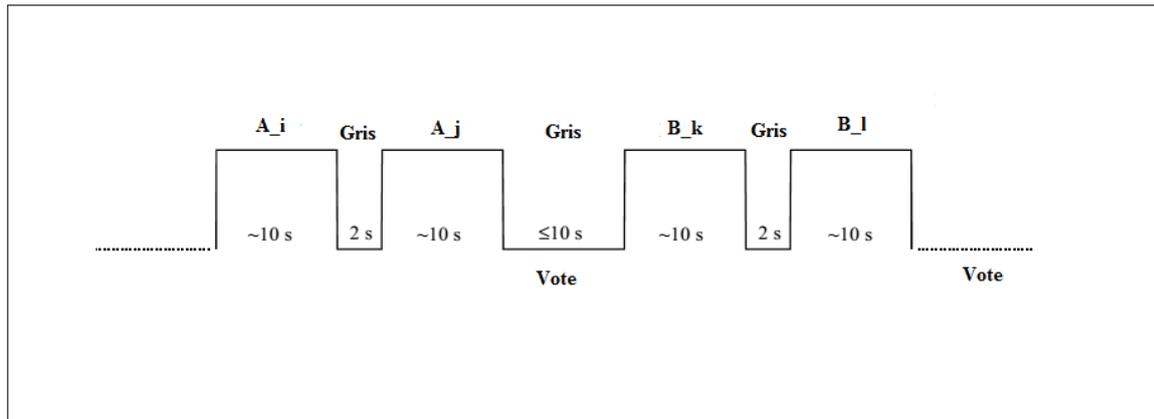


FIGURE 1.15: Méthode Comparative.

par le groupe VQEG [35].

Les résultats des tests doivent être donnés avec tout les détails de l'expérimentation. Pour chaque séquence de test la valeur moyenne des scores (Mean opinion score MOS) devrait être donnée. Elle sera utilisée pour l'évaluation des métriques objectives.

1.4.3 Bases d'images destinées à l'évaluation de qualité

Il existe plusieurs bases d'images dans la littérature dédiées à l'évaluation de la qualité perceptuelle. Dans ce paragraphe, on va rapporter celles qui sont les plus utilisées avec toutes leurs spécificités.

1.4.3.1 La base d'image LIVE [87] :

Elle est construite par l'équipe du laboratoire d'image et de vidéo à l'université Austin Texas [28]. Elle est constituée de 29 images de référence, qui ont permis de générer 779 images dégradées. Les images de référence sont montrées dans la Figure 1.16. Elle utilise cinq différents



FIGURE 1.16: Images de référence de la base LIVE.

types de dégradation : dégradations dues aux compressions JPEG2000 (JP2K) et JPEG, bruit blanc Gaussien (WN), flou Gaussien (Gblur) et la dégradation due à la transmission dans les canaux Rayleigh (FF). Pour chaque image, un score subjectif DMOS est associé. Il représente la qualité de l'image dans un intervalle de 0 à 100 où 0 représente une mauvaise qualité et 100 représente une bonne qualité.

1.4.3.2 La base Tampere TID :

Pour TID il existe deux bases, celle de 2008 [67] et celle de 2013 [66]. La seule différence entre les deux bases c'est les niveaux des dégradations utilisées. Elles ont été constituées à partir de 25 images de référence voir la Figure 1.17 . Avec 24 types de distorsions qui ont 4 (5) niveaux respectivement. Au total, il y a 1700 images dégradées pour la TID2008 et 3000 pour la TID2013. Les scores sont fournis en terme de MOS dans un intervalle de 0 à 9. La valeur la plus élevée de MOS (0 - minimal, 9 - maximal) correspond à une valeur de qualité visuelle supérieure.

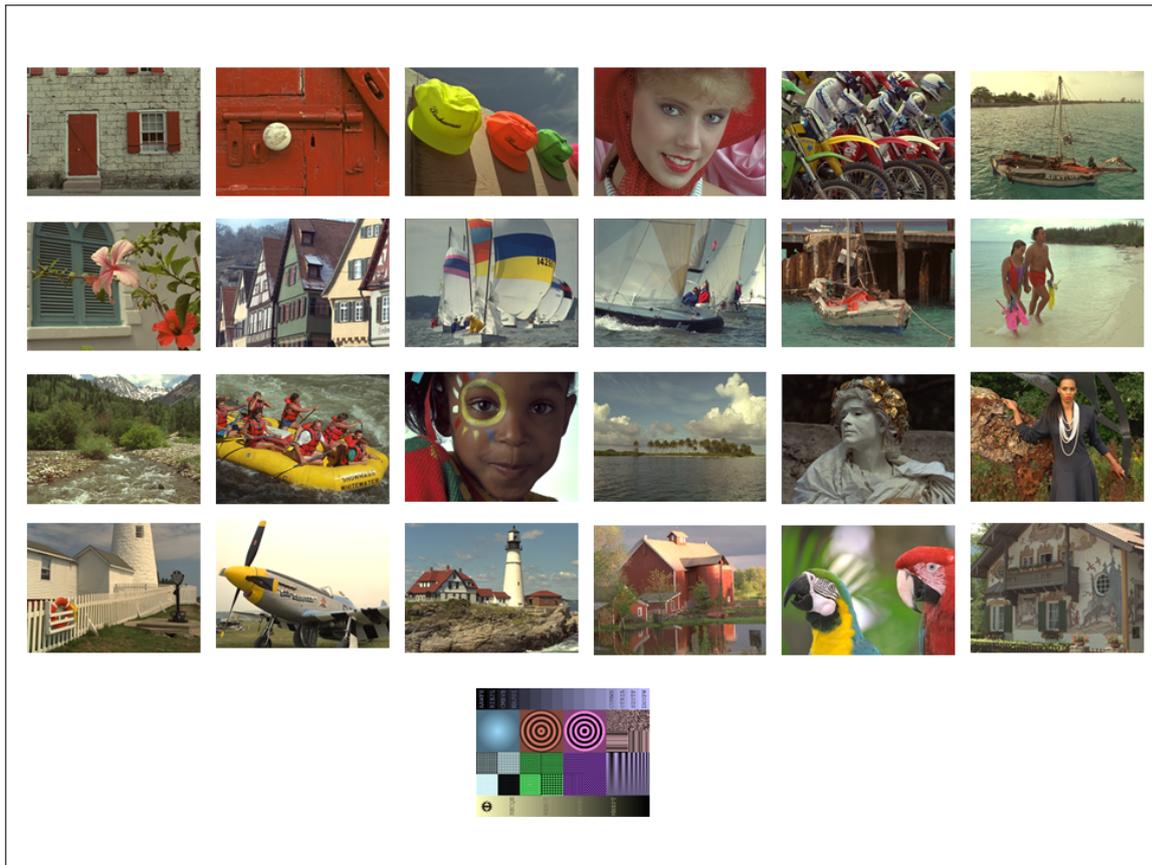


FIGURE 1.17: Images de référence de la base TID.

1.4.3.3 La base CSIQ :

La base CSIQ [40] est composée de 30 images originales Figure 1.18 où chaque image est déformée en utilisant six types de distorsions à quatre ou cinq niveaux différents. Les distorsions utilisées dans CSIQ sont : la compression JPEG, la compression JPEG2000, la dégradation du contraste, bruit Gaussien additif et flou gaussien. Un total de 866 images déformées est généré. Les images CSIQ sont évaluées subjectivement et les notes sont présentées sous la forme de DMOS.

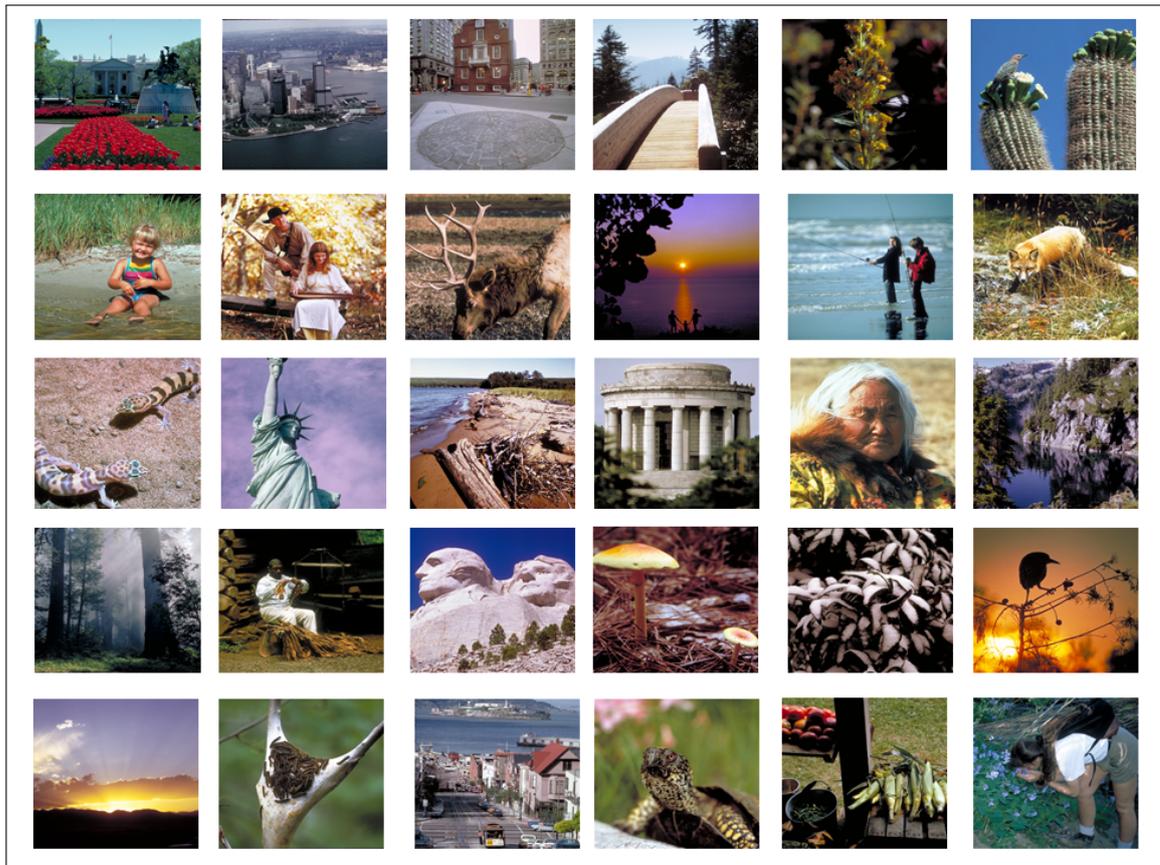


FIGURE 1.18: Images de référence de la base CSIQ.

1.4.3.4 La base IVC [64] :

Elle est générée à partir de 10 images originales voir la Figure 1.19. Les images dégradées sont obtenues en appliquant 4 algorithmes : JPEG, JPEG2000, LAR coding, Flou. Au final, 235 images ont été générées.

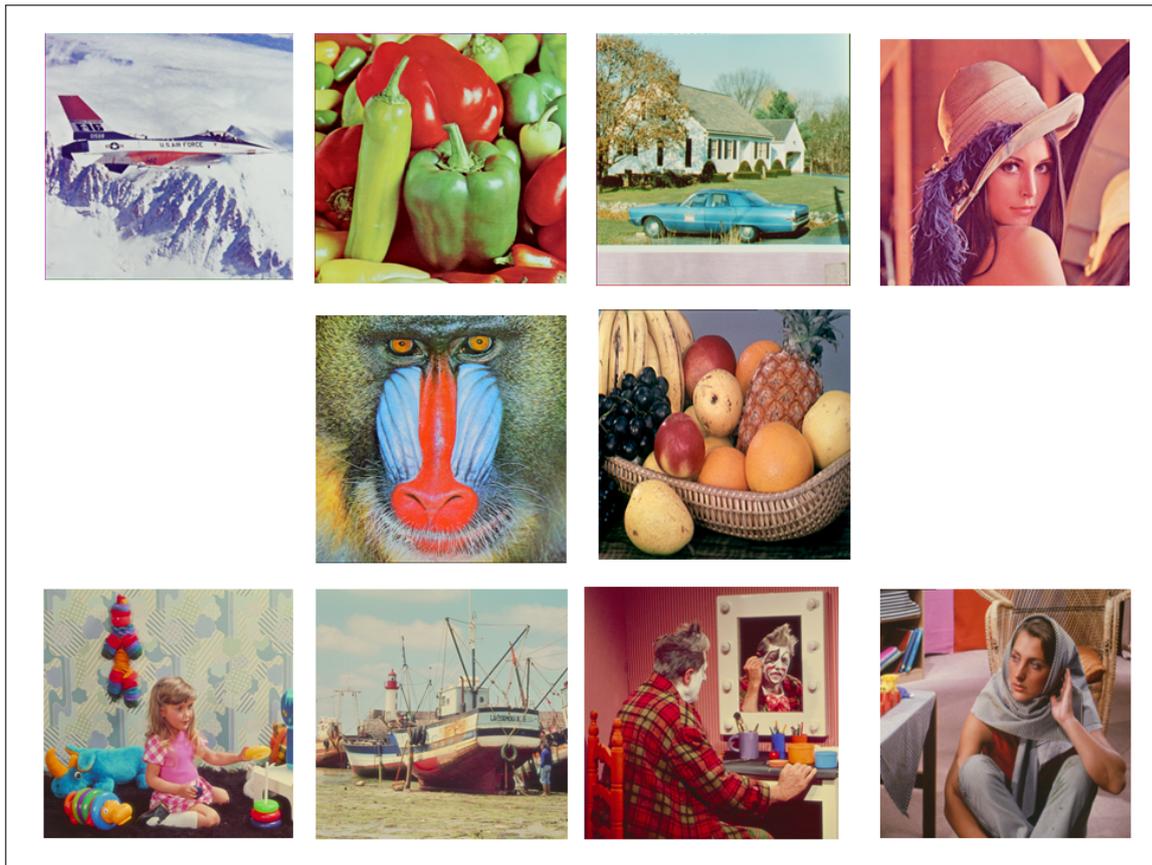


FIGURE 1.19: Images de référence de la base IVC.

Il existe également des bases distinctes qui sont destinées à l'évaluation des images stéréoscopiques (3D). On en citera quelques une par la suite.

1.4.3.5 La base LIVE 3D phase I [61] :

Elle est constituée de 20 images de références et 365 images dégradées. Il y a 80 images pour les dégradations JP2K, JPEG, WN, FF et 45 images pour la dégradation floue (Gblur). Pour chaque image, un score subjectif a été attribué sous la forme d'un DMOS. La Figure 1.20 présente les images de référence de cette base.

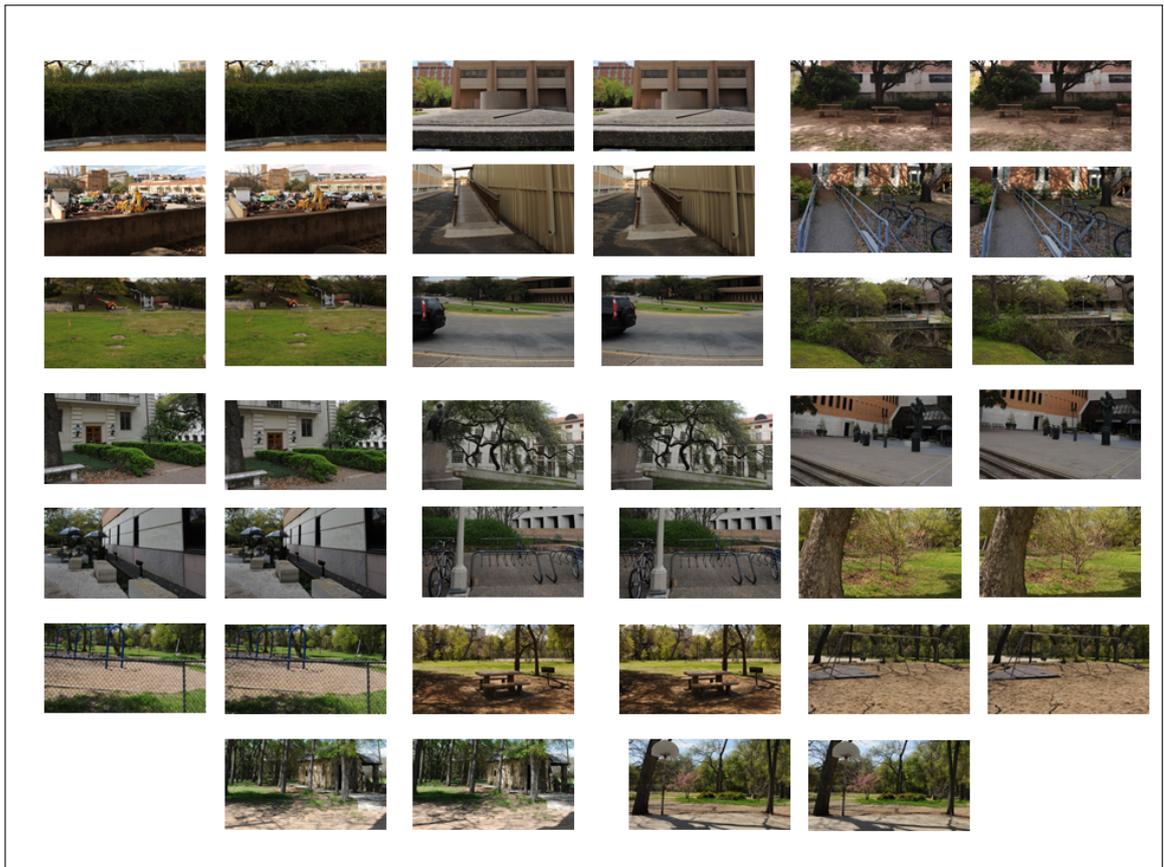


FIGURE 1.20: Images de référence de la base LIVE 3D phase I.

1.4.3.6 La base LIVE 3D phase II [11] :

Pour la phase II de la base LIVE, elle contient 8 images de références ; voir Figure 1.21 ; et 360 images dégradées. Les mêmes types de dégradation que la phase I ont été utilisées. Chaque image stéréoscopique est utilisée pour créer 3 images avec dégradations symétriques et 6 images avec dégradations asymétriques.

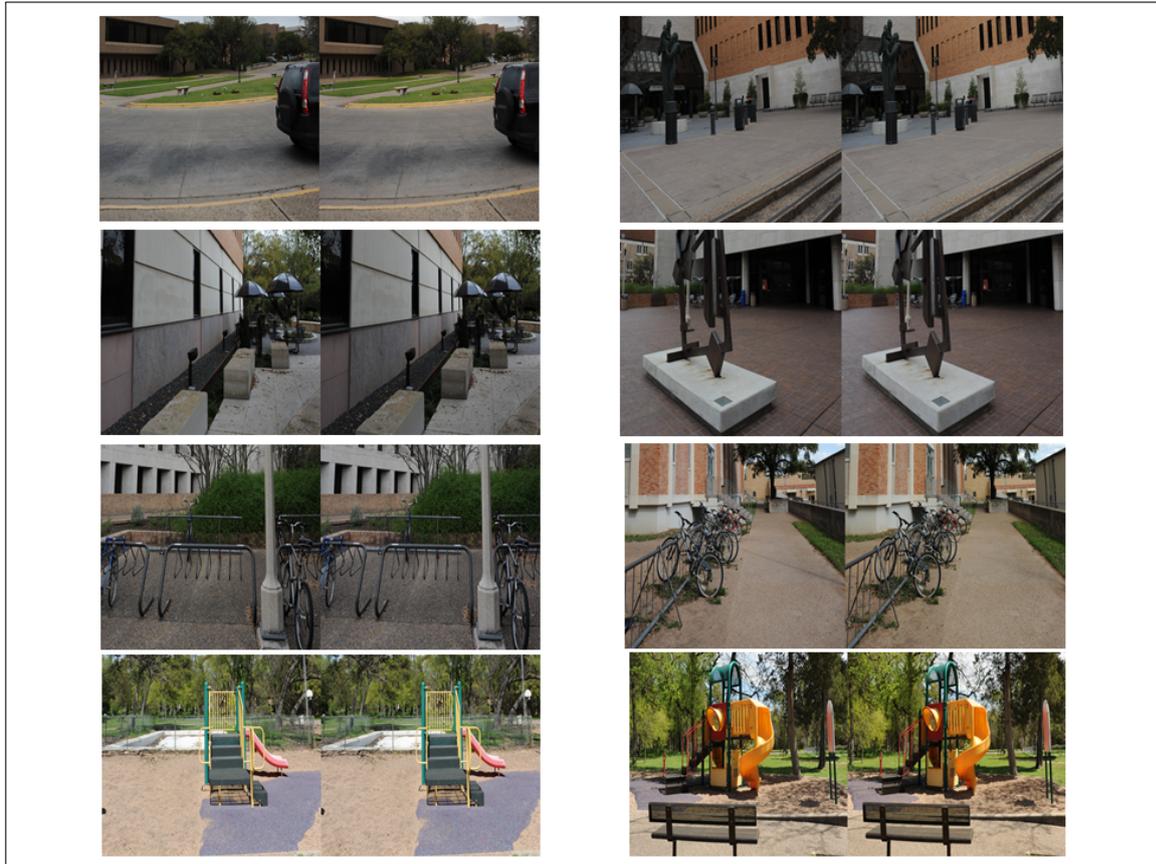


FIGURE 1.21: Images de référence de la base LIVE 3D phase II.

1.4.3.7 La base Waterloo-IVC 3D Phase I [96] :

Cette base d'images est composée de 6 images de référence et 330 images déformées [96]. Trois types de distorsions ; Bruit blanc gaussien (WN), Flou gaussien (Blur) et la dégradation due à la compression JPEG ; sont utilisés pour générer les images déformées. Pour chaque distorsion, quatre niveaux différents sont utilisés. Les paramètres de contrôle de distorsion sont mis afin d'assurer une bonne séparation perceptuelle entre les différents niveaux. Le Tableau 1.1 présente les différents groupes d'images stéréoscopiques dans la base [96] et la Figure 1.22 présente les images gauche et droite utilisées pour construire la base.

Tableau 1.1: Groupes d'images dans la base Waterloo-IVC 3D phase II

Groupe	Nombre des images	Type de distorsion
G1	6	Image originale
G2	72	Distorsion symétrique avec le même niveau
G3	72	Distorsion asymétrique sur une seule vue
G4	108	Distorsion asymétrique avec la même dégradation et différents niveaux
G5	72	Distorsion asymétrique avec dégradations et niveaux mixtes

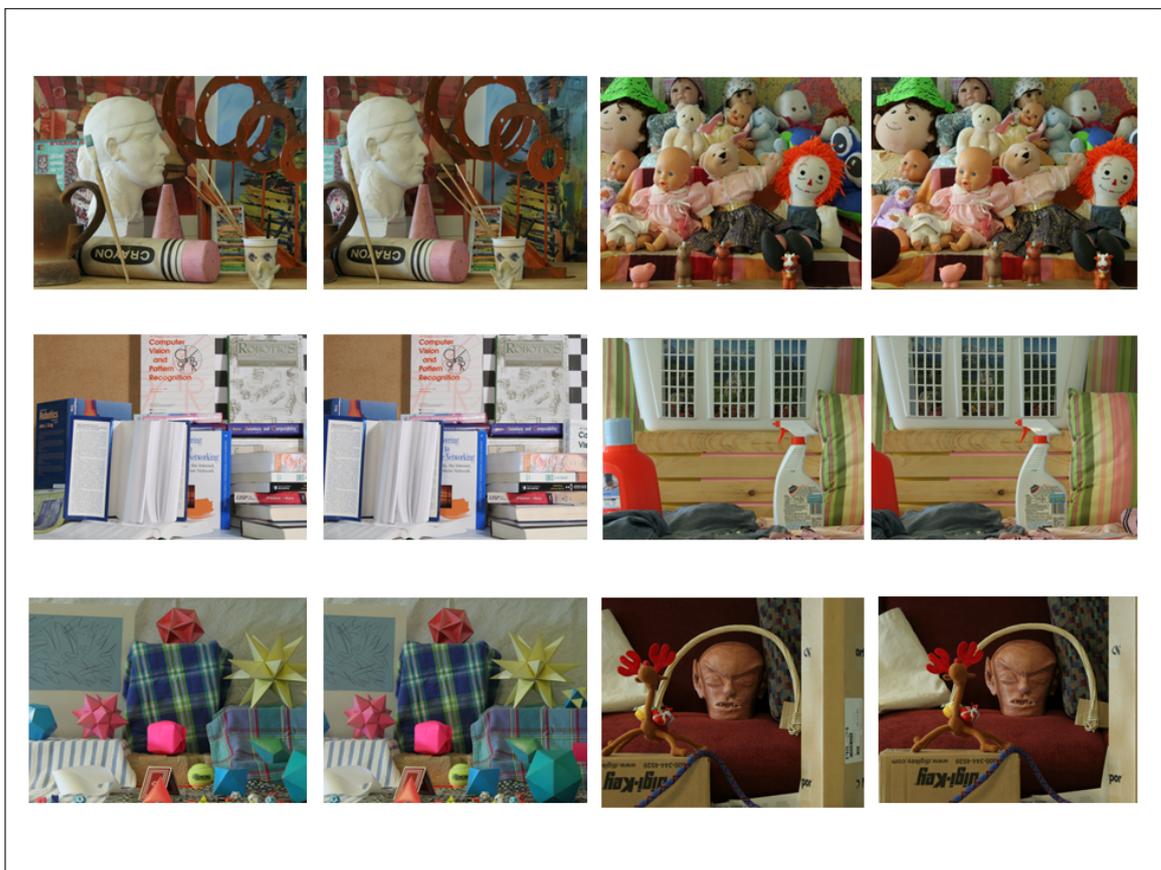


FIGURE 1.22: Images de référence de la base Waterloo-IVC 3D Phase I.

1.4.3.8 La base Waterloo-IVC 3D phase II [95] :

La phase II de la base Waterloo-IVC a été créée à partir de 10 images de référence (parfaite) Figure 1.23. Les mêmes distorsions de la phase I ont été utilisées pour dégrader chaque vue.

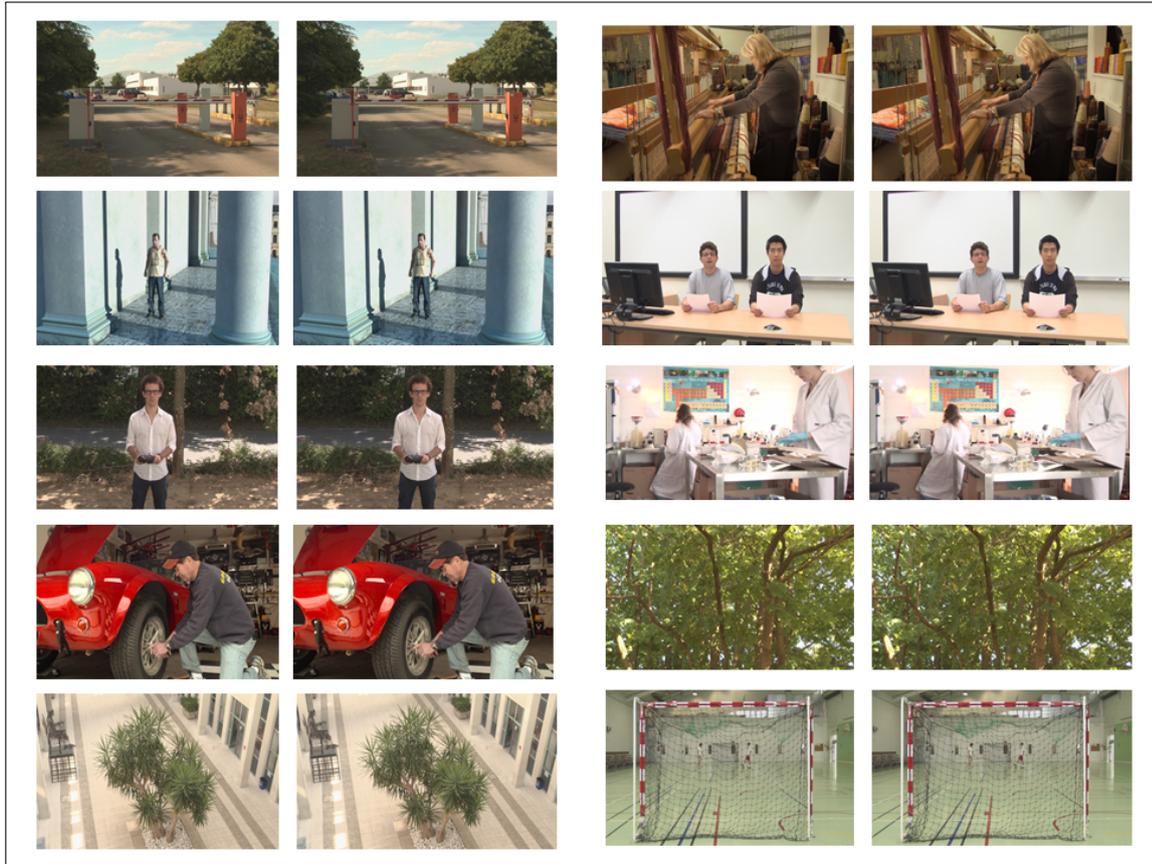


FIGURE 1.23: Images de référence de la base Waterloo-IVC 3D Phase II.

Les dégradations sont appliquées de manière symétrique et asymétrique. Au total, il y a 130 images dégradées à une seule vue et 460 images stéréoscopiques. Les scores subjectifs sont donnés en terme de score d'opinion moyen (MOS) allant de 1 à 100.

1.5 Évaluation objective et performance

L'évaluation subjective reste la meilleure manière d'estimer la qualité visuelle. Le problème pour ce genre de méthodes est le coût en terme de temps et de ressource. Cela a conduit les chercheurs à penser à d'autres méthodes d'évaluation qui sont éventuellement les métriques objectives. En ce qui concerne ce genre, l'évaluation de la qualité perceptuelle est faite automatiquement en utilisant des algorithmes. Elles utilisent pour leurs conception les caractéristiques mathématiques de l'image qui sont susceptibles d'être affectées par les artéfacts.

Les métriques d'évaluations objectives se séparent en trois groupes : les métriques avec référence MAR, les métriques avec référence réduite MRR et les métriques sans référence MSR.

Toute nouvelle métrique proposée doit être évaluée pour connaître son efficacité. Les sorties de ces dernières doivent être corrélées avec les scores subjectifs de manière prévisible et répétitive. La linéarité entre les deux scores objectifs et subjectifs n'est pas cruciale. Le plus important et déterminant pour la performance est la stabilité de la relation. Pour enlever toute non linéarité due à l'évaluation subjective ; voir la Figure 1.24 ; et pour faciliter la comparaison des scores, la performance sera estimée à l'aide d'une régression non linéaire entre l'ensemble des scores objectifs et subjectifs [35].

Après avoir effectué la transformation linéaire, on obtient les scores d'opinion moyens MOS_p qui sont les scores prédits. En utilisant les paramètres cités après, la performance de la métriques peut être évaluée.

- **Le coefficient de corrélation linéaire de Pearson (PLCC) :** C'est la covariance entre les scores subjectifs MOS et objectifs MOS_p .

$$(1.3) \quad PLCC = \frac{\sum_{i=1}^N (MOS_{p_i} - \bar{MOS}_p)(MOS_i - \bar{MOS})}{\sqrt{(\sum_{i=1}^N (MOS_{p_i} - \bar{MOS}_p)^2)(\sum_{i=1}^N (MOS_i - \bar{MOS})^2)}}$$

où N est le nombre d'image utilisé. MOS_p , MOS , \bar{MOS}_p et \bar{MOS} sont le score objectif, le score subjectif et leurs valeurs moyennes respectivement.

- **Le coefficient de corrélation de Spearman (SROCC) :** il calcule l'ordre dans lequel les valeurs des deux vecteurs MOS_p et MOS apparaissent.

$$SROCC = 1 - \frac{6}{N(N^2 - 1)} \sum_{i=1}^N d_i^2$$

où d_i est la différence entre les rangs du score objectif et du score subjectif de l'image testée.

- **La racine de l'erreur moyenne quadratique (RMSE) :** C'est la racine de la moyenne carrés des écarts entre les scores subjectifs MOS et objectifs MOS_p .

$$(1.4) \quad RMSE = \sqrt{\sum_{i=1}^N (MOS_{p_i} - MOS_i)^2}$$

MOS_p et MOS sont les scores objectifs après fusion et les scores subjectifs.

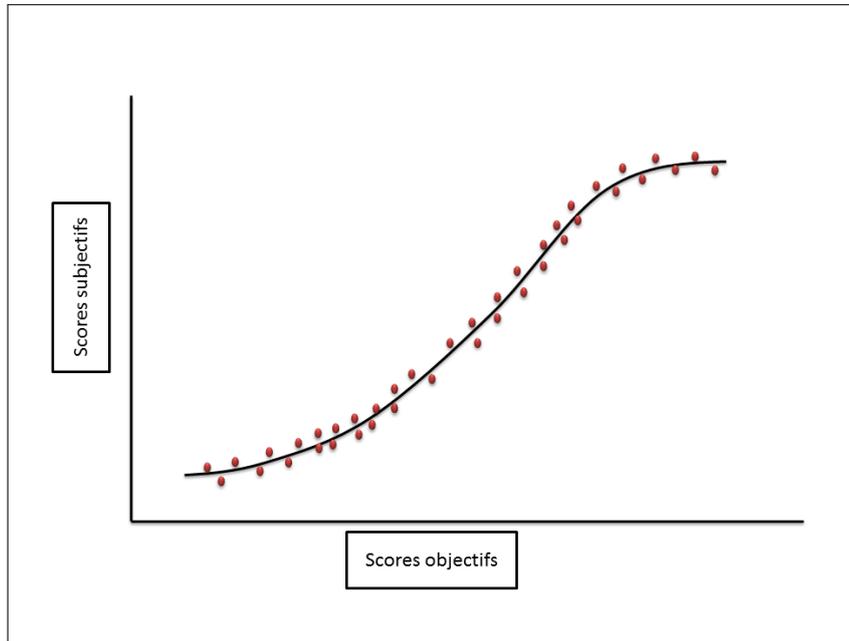


FIGURE 1.24: Exemple de relation entre les scores objectifs et subjectifs.

1.6 Les applications de la qualité d'image

L'intérêt envers le développement des métriques de qualité ne cesse pas de croître. Cela est due aux progrès dans le domaine de l'information et les nouveaux enjeux qui les accompagnent. Dans ce paragraphe, on donne un aperçu sur les potentielles applications des métriques de qualité.

Une des applications les plus simples des métriques d'évaluation de qualité (MEQ) peut être la comparaison des algorithmes et systèmes de traitement. Prenons l'exemple des images dégradées par le bruit (ou par le flou). Après la restauration, une MEQ peut être utile pour choisir l'image qui a la meilleure qualité visuelle.

À côté des scores qu'elles produisent, les MEQ peuvent aussi générer des cartes de qualité. Une carte de qualité a une grande importance, puisqu'elle représente les endroits les plus affectés par la distorsion et elle varie en fonction de sa sévérité. Cette dernière peut être utilisée comme cartographie pour les algorithmes de denoising [31].

Les mesures de qualité peuvent être utilisées pour optimiser et déterminer les paramètres des algorithmes de traitement. Par exemple, la fixation automatique des paramètres des standards de compression en se basant sur la qualité perceptuelle comme critère.

La qualité visuelle est particulièrement souhaitable dans les applications de communication visuelle sur les réseaux. Les images transmises subissent des dégradations de la qualité, cela est due à la compression et aux canaux de transmission. Il est impératif pour les fournisseurs de service réseau de surveiller les dégradations de qualité en temps réel, afin d'optimiser les allocations des ressources du réseau et satisfaire au maximum les attentes des utilisateurs. Par conséquent, l'évaluation de qualité d'image perceptuelle peut jouer le rôle de contrôle et de surveillance en temps réel.

Dans la plupart des applications citées précédemment, le résultat final est une image. En outre, les mesures de qualité peuvent également être étendues à d'autres applications où le résultat sera une interprétation ou une classification. La reconnaissance d'image est un tel exemple où la qualité des images est un facteur crucial qui affecte directement la précision de l'algorithme de reconnaissance. En biométrie, l'objectif est de reconnaître ou de vérifier l'identité de la personne sur la base d'une ou plusieurs caractéristiques physiologiques uniques de l'être humain. En pratique, le processus d'acquisition n'est pas toujours parfait, et cela peut produire des images déformées ou partiellement défectueuses. Dans ce cas, il serait utile de connaître le niveau de dégradations qui sert à définir la précision de la reconnaissance à appliquer. Ou bien pour relancer un prétraitement afin d'améliorer la qualité de l'image et aider le processus de reconnaissance [97].

Le domaine de la médecine est un domaine qui a connu des progrès rapides notamment dans l'imagerie médicale. Le nombre d'images médicales acquises chaque jour est considérable, ce qui nécessite un espace de stockage important. Pour faciliter l'enregistrement, la transmission

et la récupération de ces images la compression est la seule solution, sachant les dégradations qu'elle peut introduire. Certes, elle représente un moyen puissant pour réduire la taille des données. Mais, il y a un grand risque de perdre des informations de diagnostics importantes. Il est donc impératif d'avoir des métriques de qualité qui peuvent aider à maximiser le niveau de compression sans affecter les diagnostics médicaux [97].

1.7 Conclusion

Ce chapitre était une introduction à la qualité d'image, dans lequel on a pu dans un premier temps donner une définition générale de l'évaluation de la qualité. Ensuite, nous avons traité les dégradations les plus connues et souligner leurs impacts sur l'image et les nuisances qui en découlent. Des images qui montrent les artefacts et distorsions comme le bruit, le flou et les effets de compression ont été rapportées pour mieux éclairer leurs effets.

Par la suite, l'évaluation subjective a été présentée. Une description générale du mode de fonctionnement du système visuel humain SVH a été donnée. Le rôle de chaque organe de la chaîne de traitement a été décrit brièvement.

Comme l'évaluation subjective est importante, des protocoles d'évaluations qui assurent le déroulement des expérimentations sont obligatoires. Les protocoles mis en place par la communauté scientifique ont été présentés. La dernière partie de l'évaluation subjective est la construction des bases. Pour cela, les bases les plus utilisées dans la littérature ont été présentées.

Les deux dernières sections ont été consacrées à l'évaluation objective. Une définition de cette dernière et les méthodes d'évaluation de performance ont été données. Et enfin, des exemples sur l'utilité et les applications des métriques objectives sont adressés.

MÉTRIQUES DÉVELOPPÉES POUR L'ÉVALUATION DE LA QUALITÉ D'IMAGE : ÉTAT DE L'ART

Sommaire

2.1	Introduction	34
2.2	Métriques avec référence (MAR-EQI)	34
2.3	Métriques avec référence réduite (MRR-EQI)	35
2.4	Métriques sans référence (MSR-EQI)	36
2.4.1	Métriques basées sur les statistiques de la scène naturelle (NSS)	37
2.4.2	Métriques basées sur les réseaux de neurones	40
2.4.3	Métriques basées sur les dictionnaires (codebooks)	41
2.4.4	Limitations des métriques sans référence existantes	43
2.5	Conclusion	43

2.1 Introduction

L'évaluation de la qualité d'image est un domaine qui a été créé pour répondre aux différents enjeux dans lesquelles l'information visuelle est utilisée. En commençant par l'acquisition, l'image peut subir différentes distorsions qui dégradent sa qualité. En revanche, dans toute la chaîne du post-traitement ou du pré-traitement la qualité est mise en question. En effet, il est nécessaire d'avoir des mesures de qualité, capables d'évaluer le contenu visuel en permanence. De ce fait, les chercheurs ont beaucoup investi ces dernières années dans ce domaine, afin d'élaborer des métriques de qualité performantes et qui peuvent répondre aux différentes problématiques. Les travaux de qualité d'images peuvent être classés en trois groupes : qualité d'image avec référence, qualité d'image avec référence réduite et qualité d'images sans référence. Dans les paragraphes suivants, un état de l'art de ce qui a été proposé dans la littérature est donné.

2.2 Métriques avec référence (MAR-EQI)

Dans les métriques avec référence MAR-EQI, l'information complète de l'image de référence est indispensable pour évaluer la qualité de l'image dégradée. Les approches avec référence sont basées sur la mesure de disparité entre les pixels de l'image de référence et l'image déformée. Parmi les métriques les plus connues on a :

- **La moyenne de l'erreur quadratique (MSE) :**

$$(2.1) \quad MSE = \frac{1}{M \times N} \sum_{m=1}^M \sum_{n=1}^N (I(m,n) - \hat{I}(m,n))^2$$

avec I et \hat{I} les images originale et dégradée respectivement, (M, N) sont les tailles de l'image.

- **le rapport signal sur bruit (PSNR) :**

$$(2.2) \quad PSNR = 10 \log_{10} \left(\frac{I_{max}^2}{MSE} \right)$$

avec (I_{max}) est la valeur maximal des niveaux de gris.

Cependant ces deux métriques ont montré leurs incapacité d'évaluation. En général, leurs scores ne correspondent pas aux évaluations subjectives.

Par ailleurs, plusieurs métriques plus performantes ont été proposées. L'information structure a été exploitée pour l'évaluation. Une image de bonne qualité possède une structure similaire de l'image originale. En effet, la luminance, la structure et le contraste sont, de ce fait utilisés pour estimer la qualité.

- **L'indice de similarité structurelle (SSIM) [101] :** La mesure SSIM est basée sur la mesure de similarité structurelle entre l'image originale et dégradée. L'image est décomposée en blocs. Les composantes luminance, contraste et structure sont calculées sur chaque

blocs des deux images à comparer. Elles sont données comme suit :

La luminance :

$$(2.3) \quad l(I, \hat{I}) = \frac{2\mu_I\mu_{\hat{I}} + (C_1L)^2}{\mu_I^2 + \mu_{\hat{I}}^2 + (C_1L)^2}$$

avec C_1 une constante pour éviter l'instabilité quand le dénominateur est proche de zéro, L est le niveau de gris de l'image $L = 255$, I et \hat{I} représentent les images originale et dégradée respectivement.

La contraste :

$$(2.4) \quad c(I, \hat{I}) = \frac{2\sigma_I\sigma_{\hat{I}} + (C_2L)^2}{\sigma_I^2 + \sigma_{\hat{I}}^2 + (C_2L)^2}$$

avec C_2 une constante de faible valeur.

La structure :

$$(2.5) \quad s(I, \hat{I}) = \frac{2\sigma_{I\hat{I}} + (C_2L)^2}{2\sigma_I^2\sigma_{\hat{I}}^2 + (C_2L)^2}$$

où les μ et les σ sont les moyennes et les covariances respectivement.

L'indice de similarité globale est le résultat de produits des trois composantes :

$$(2.6) \quad SSIM(I, \hat{I}) = l(I, \hat{I}).c(I, \hat{I}).s(I, \hat{I})$$

avec l , c , s la luminance, le contraste et la structure respectivement.

À partir de la métrique SSIM de base différentes versions ont été proposées. On peut citer SSIM multi-échelle (MS-SSIM) [103], SSIM d'ondelette complexe (CW-SSIM)[74], SSIM basée sur la force des contours (ESSIM) [117], l'indice de qualité d'image universelle (UQI) [98] et l'indice de similarité des caractéristiques (FSIM) [116].

En plus des algorithmes basés sur la structure des images, il existe d'autres travaux qui ont intégré les caractéristiques du système visuel humain (SVH). Pour cela, le modèle du cortex visuel primaire (V1) a été employé. Les métriques rapport signal sur bruit visuel (VSNR) [8] et celle des distorsions les plus apparentes (MAD) [41] représentent bien l'exploitation des caractéristiques du système visuel humain (SVH).

2.3 Métriques avec référence réduite (MRR-EQI)

En revanche, les métriques avec référence réduite n'ont pas besoin de l'intégralité de l'image de référence. En effet, il suffit d'avoir quelques caractéristiques représentatives pour évaluer la

qualité. Une chaîne auxiliaire est utilisée pour l'envoi de ces caractéristiques. Du côté récepteur, des caractéristiques identiques sont extraites afin d'être comparées avec celles transmises. La Figure 2.1 représente le mode de fonctionnement des métriques avec référence réduite MRR-EQI. En général les MRR-EQI sont basées sur des modèles qui utilisent des propriétés statistiques de

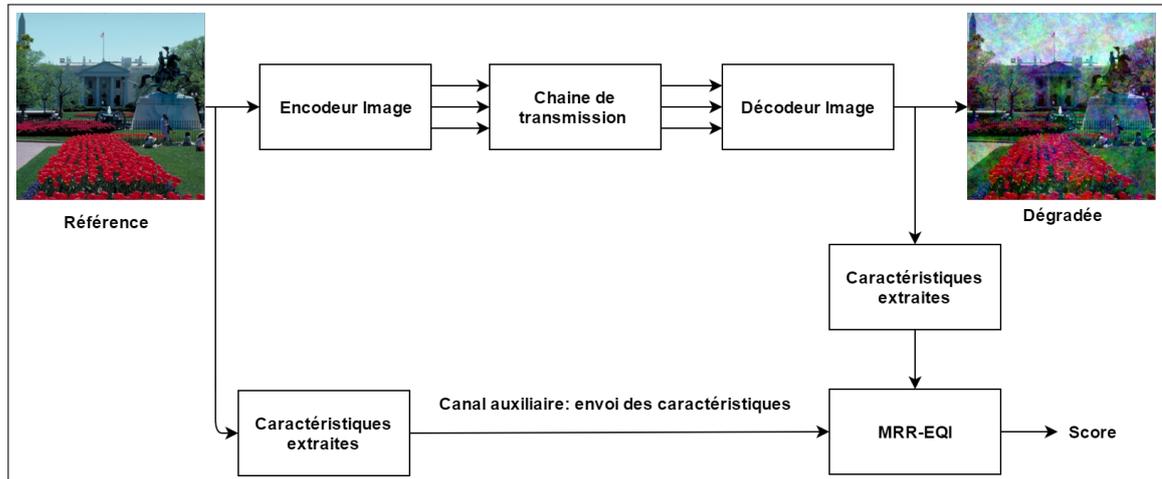


FIGURE 2.1: Le principe des métriques avec référence réduite MRR-EQI

l'image dans les domaines de transformation. Parmi les statistiques les plus utilisées dans les MRR-EQI, il y a la transformée d'ondelette discrète (DWT) [45]. Autres chercheurs ont choisi d'utiliser les caractéristiques basées sur le type de distorsion affectant l'image. Par exemple les standards de compression [14, 39].

Les deux types de métriques, MAR-EQI et MRR-EQI ont eu des résultats satisfaisants. Néanmoins, le principal inconvénient de celles-ci, est la dépendance totale ou partielle de l'image de référence. Pour cela, un troisième type sans référence a été proposé dans la littérature.

2.4 Métriques sans référence (MSR-EQI)

Les métriques sans référence MSR-EQI (blind) représentent une partie importante de l'évaluation de la qualité d'image. On peut dire que le recours à l'évaluation sans référence est problématique de part sa difficulté. Toutefois, son usage apporte des avantages qui ne sont pas négligeables. Celle-ci est difficile en raison de l'absence de l'image de référence et de certaines caractéristiques. En revanche, l'évaluation sans référence apporte des avantages en termes d'efficacité et de praticité. En effet, celle-ci est adaptée à des applications réelles où nous n'avons pas forcément accès à l'image de référence. Dans l'évaluation sans référence on peut distinguer deux catégories de métriques ; les métriques qui sont basées sur le type de dégradations. Dans cette catégorie, le type de distorsion est connu à l'avance. Pour la seconde, la dégradation qui affecte l'image n'est pas prise en compte.

Les métriques avec spécification de distorsion traitent souvent des dégradations comme le flou, le bruit, les dégradations d'effet de bord (de bloc) et les dégradations dues à la compression (JPEG, JPEG2000). Les travaux [10, 16, 22, 23, 38, 54, 60, 79, 93, 109] ont bien traité l'évaluation des dégradations flou et bruit.

Ces dernières années, le domaine de l'image s'est nettement développé grâce à de nouvelles acquisitions visant l'amélioration de la qualité des supports visuels. Cette meilleure qualité est traduite par des volumes de stockage important. Pour faciliter leurs acheminement dans les réseaux, les chercheurs ont été obligés d'utiliser les processus de compression. Néanmoins, le principal inconvénient avec la compression réside dans le fait qu'elle est appliquée dans des domaines différents de ceux de l'image. Cela introduit de multiples dégradations qui influencent la qualité de l'image reçue. Parmi les standard les plus utilisés, on a la compression JPEG et la compression JPEG2000. La compression JPEG introduit un effet de bloc ou effet de bord à l'image. Différents travaux ont été proposés pour évaluer la distorsion engendrée par la compression JPEG [2, 17, 24–26, 100, 106, 108].

D'une autre part, la compression JPEG2000 introduit deux types de distorsion à l'image ; le flou et le ringing. Le ringing représente des oscillations autour des contours. Malgré la difficulté de la procédure de l'évaluation, il y avait de nombreuses métriques qui ont été proposées dans la littérature. Parmi ces dernières, on peut citer [46, 48, 54, 75, 76, 76, 94, 114, 115].

Malgré le nombre de métriques avec spécification de distorsion proposées dans la littérature, elles restent toujours loin de répondre aux différentes contraintes notamment dans le cas où plusieurs dégradations sont présentes dans la même image. De ce fait, il faut mettre en œuvre des métriques qui sont capables d'évaluer la qualité sans avoir recours aux distorsions affectant l'image.

Pour ce faire, les chercheurs ont proposé de nouvelles mesures dites génériques. Elles peuvent être classées selon le modèle qu'elles utilisent en trois groupes. Les mesures qui utilisent les caractéristiques de la scène naturelle (natural scene statistics NSS), les mesures qui ont recours aux dictionnaires (codebooks) et les mesures qui adoptent les réseaux de neurones. Il faut noter que dans une mesure de qualité, on peut trouver un mélange entre les trois classes. Par exemple, une métrique qui utilise les NSS et les réseaux de neurone en même temps.

2.4.1 Métriques basées sur les statistiques de la scène naturelle (NSS)

Dans [70] Ruderman a étudié l'importance des statistiques de la scène naturelle. Il a montré que les images naturelles possèdent des statistiques spécifiques. Ruderman a pu montrer que ces statistiques ont une forme régulière. La variance locale normalisée du contraste et les statistiques de la relation entre les pixels voisins peuvent avoir une forme Gaussienne.

En se basant sur cette étude, plusieurs travaux basés sur les statistiques NSS ont été proposés dans littérature. Il y a deux types d'utilisation de ces dernières ; l'extraction direct de l'image (sur les pixels) ou bien l'extraction dans le domaine transformé de l'image. Le principe général des

métriques sans référence qui utilisent les NSS est montré dans la Figure 2.2.

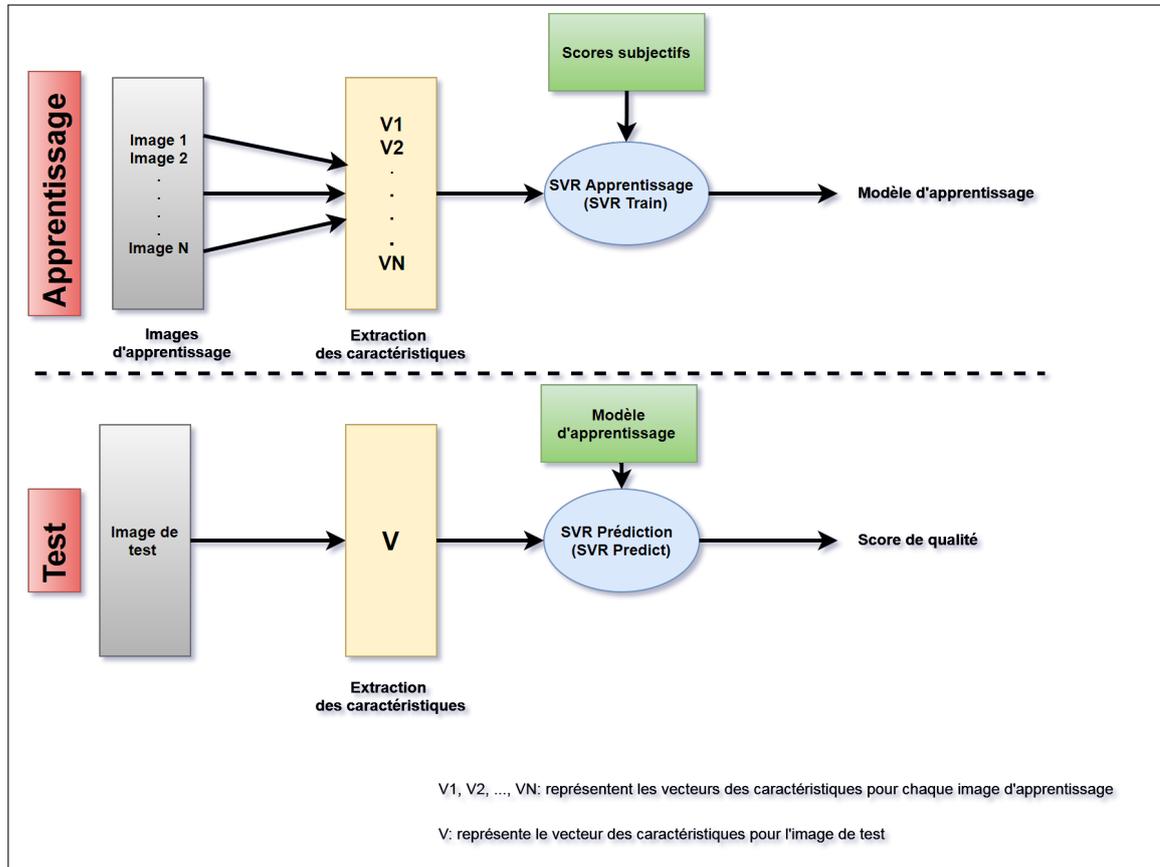


FIGURE 2.2: Principe général des métriques basées sur les NSS.

La métrique proposée dans [60] est une métrique basée sur la décomposition en ondelette. En utilisant la décomposition pyramidale orientée, l'image a été décomposée en 12 sous-bandes (deux échelles et six orientations). Ensuite, les statistiques de la scène naturelle NSS sont extraites de chaque sous-bandes. Les NSS sont paramétrées en utilisant une distribution Gaussienne généralisée pour produire au total 24 caractéristiques (2 échelles \times 6 orientations \times 2 paramètres). Un autre ensemble de caractéristiques qui représente la dépendance entre les coefficients des sous-bandes à travers les échelles et les orientations a été ajouté pour caractériser les différentes distorsions. Enfin, les caractéristiques indiquant les corrélations entre les échelles sont également calculées. Toutes ces caractéristiques ont été utilisées afin de représenter l'allure des statistiques des images naturelles. Pour estimer la qualité de l'image un processus de classification/régression a été mis en place.

Zhang et al ont proposé C-DIVINE [119] une nouvelle version de la métrique proposée dans [60]. La métrique applique tout d'abord une décomposition pyramidale complexe orientée pour chaque image dégradée. Le but de cette décomposition est d'obtenir des caractéristiques complexes multi-

échelles avec différentes orientations. Comme dans [60], les coefficients sont ensuite normalisés et modélisés en utilisant une distribution gaussienne. C-DIIVINE mesure aussi l'amplitude et la phase des caractéristiques statistiques utilisées dans DIIVINE en utilisant la métrique SSIM. Toutes les caractéristiques des NSS ont été combinées afin d'être utilisées en entrée pour la machine à vecteur de support SVR.

La transformée en cosinus discrète DCT a été utilisée dans la métrique proposée dans [72]. BLIINDS (nom de la métrique) est basée sur un seul étage de décomposition DCT. La DCT est calculée premièrement sur des blocs de taille 17×17 centrés à chaque pixel de l'image. Les statistiques sont ensuite extraites à partir de deux échelles. Elles représentent les informations de contraste et de structure dans l'image. Les auteurs ont calculé par la suite la valeur moyenne locale du contraste sur tous les blocs DCT. La deuxième caractéristique a été déterminée en calculant le kurtosis de chaque bloc DCT de l'image, puis, en faisant la moyenne des dix pour cent les plus bas des valeurs calculées. Deux autres caractéristiques ont été calculées, la variance de l'entropie et la valeur maximale de l'entropie des coefficients DCT pour les quatre directions (0° , 45° , 90° et 135°). Une fois le calcul des paramètres terminé, la distribution gaussienne multivariée est utilisée pour modéliser ces caractéristiques. A l'étape d'apprentissage, un modèle a été obtenu en utilisant les caractéristiques et les DMOS correspondants pour chaque image. Le modèle est ensuite utilisé pour obtenir le score final de la qualité.

L'algorithme DESIQUE [118] utilise les caractéristiques de [56] et les élargit en utilisant d'autres propriétés qui sont extraites dans le domaine fréquentiel. Il est basé sur l'observation suivante : la dérivée de certaines propriétés statistiques dans les domaines spatial et fréquentiel varieront en présence de la distorsion. De plus, en étudiant ces variations, la qualité de l'image peut être évaluée. Les coefficients locaux de luminance normalisés et leurs produits sont calculés et modélisés par une distribution Gaussienne généralisée GDD. Le deuxième ensemble de caractéristiques a été développé en modélisant la relation de voisinage des coefficients de luminance en utilisant les statistiques de log-dérivées.

Dans le domaine fréquentiel, l'image est d'abord décomposée en utilisant un filtre log-Gabor sur deux orientations. Ensuite, les mêmes statistiques du domaine spatial ont été calculées. Comme les images sont naturellement multi-échelle et les distorsions changent la structure de l'image à l'échelle, les statistiques sont extraites sur deux échelles. Pour produire le score de qualité, le SVR a été utilisé.

La métrique nommée Brisque [56] représente l'une des métriques les plus importantes qui sont basées sur les NSS. Les statistiques sont calculées sur le domaine spatial de l'image. Les coefficients locaux de luminance normalisés et leurs produits sont utilisés pour prédire la qualité de l'image. Ces caractéristiques sont sélectionnées à partir des observations qu'elles possèdent des propriétés statistiques qui varient en présence des dégradations. Les signes des coefficients adjacents montrent également une structure régulière qui peut être influencée par les dégradations. La distribution des coefficients de luminance a été modélisée par une GGD.

La variance et le paramètre de forme sont ensuite calculés. Les produits des coefficients adjacents ont été calculés selon quatre directions et modélisés par une GGD asymétrique. À la fin, 36 paramètres ont été obtenus sur deux échelles. Ils sont ensuite utilisés comme entrée pour le SVR afin de réaliser le mappage (passage) de l'espace de fonctionnalité à la qualité de l'image.

Dans [90], de simples caractéristiques ont été utilisées pour estimer la qualité de l'image. Les caractéristiques dérivées des statistiques NSS, des caractéristiques de texture et d'estimation du flou/bruit ont été calculées. Des analyses approfondies ont pu montrer que l'amplitude et la phase des réponses des filtres hautes fréquences contiennent plus d'informations et elles peuvent capturer des artefacts de compression. Les caractéristiques ont été combinées en trois groupes distincts, répartition marginale des coefficients d'ondelettes, répartition des coefficients d'ondelettes à travers les différentes échelles et statistiques de flou/bruit. Dans la phase d'apprentissage, l'analyse en composante principale ACP a été d'abord faite pour chaque groupe afin de réduire la dimension des caractéristiques. Le résultat de l'ACP est ensuite utilisé pour former un modèle de régression en utilisant le SVR.

2.4.2 Métriques basées sur les réseaux de neurones

Les métriques basées sur les réseaux de neurones, Figure 2.3, suivent la même ligne que celles basées sur les NSS. Au départ, un certain nombre de caractéristiques représentant les spécificités de l'image sont calculées. Ensuite, elles sont utilisées comme entrée pour les premières couches du réseau. Le réseau de neurone prend ensuite le relais à l'aide de ces couches cachées pour estimer la qualité qui va être donnée par la couche de sortie. Certaines caractéristiques

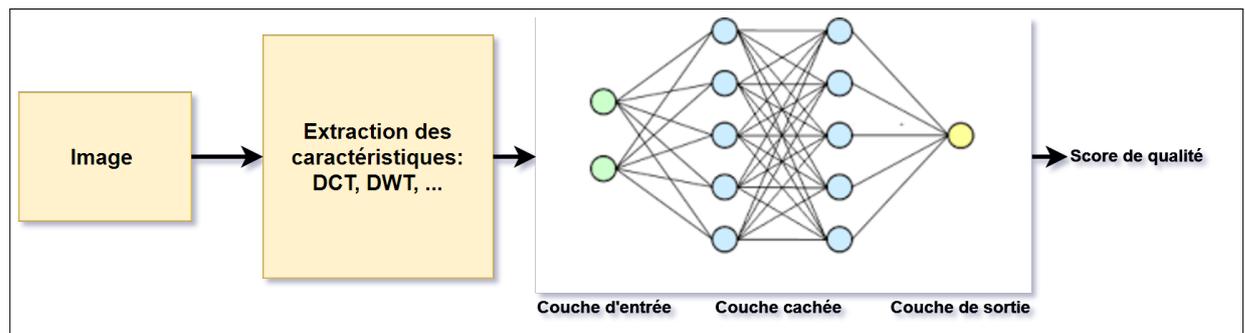


FIGURE 2.3: Principe général des métriques basées sur les réseaux de neurones.

de l'image possédant des informations sur le contenu perceptuel ont été utilisés dans [42]. Les caractéristiques sont conçues en fonction des trois principaux attributs de l'image : congruence de phase, l'entropie locales et le gradient. La congruence de phase contient des informations importantes qui se produisent dans des endroits spatiaux où les composantes de la transformée de Fourier sont maximales. L'entropie a été utilisée pour mesurer l'anisotropie de l'image.

Le gradient a été considéré dans la métrique pour estimer les variations de luminance de structure et de texture dans l'image.

Après l'obtention des caractéristiques, les réseaux généraux de régression (GRNN) ont été utilisés pour calculer la qualité. Ils sont constitués de quatre couches : la couche d'entrée, la couche de motif, la couche de sommation et la couche de sortie. Le nombre de couches est égal au nombre de paramètres. Chaque unité dans la couche de motif représente un modèle d'apprentissage. Chaque nœud dans la couche de motif est connecté à chacun des deux nœuds dans la couche de sommation. La dernière couche produit le score de qualité à partir des couches de sommation. Kang et al [37] ont développé une métrique MSR-EQI basée sur les réseaux de neurones convolutifs. La métrique combine les caractéristiques d'apprentissage et la régression comme un processus d'optimisation complet. Les blocs de l'image sont utilisés comme entrées pour le réseau de neurones qui est constitué d'une couche convolutionnelle, avec un étage de mise en commun (pooling) max et min, deux couches totalement connectées et un nœud de sortie. En fonction de l'apprentissage, le réseau peut donner le score global de la qualité.

Dans [91] une métrique basée sur les réseaux de neurones a été proposée. Elle estime la qualité de l'image à partir des caractéristiques pertinentes extraites. Pour cela, ils ont utilisé les mêmes caractéristiques que dans [90]. Le score de qualité a été calculé en utilisant un modèle de réseau multi-couches. Il transforme les caractéristiques par régression à un seul score de qualité. Le modèle est constitué de deux composants : le premier composant est un processus gaussien qui donne le score de qualité finale à partir d'une activation du réseau neuronal d'apprentissage. Le deuxième composant est un réseau neuronal dont le but est de fournir une représentation informative des caractéristiques pour l'évaluation de la qualité. Ils ont adopté le réseaux de neurones d'apprentissage profond (deep belief network (DBN)) pour faire cette architecture.

2.4.3 Métriques basées sur les dictionnaires (codebooks)

La notion des dictionnaires (codebook) a été utilisée dans les MSR-EQI. Les dictionnaires reposent sur le principe suivant ; le signal image peut être décomposé en une combinaison linéaire de quelques signaux. Ceux-ci sont utilisés pour former un dictionnaire. Ce dernier comporte d'importantes propriétés de l'image qui seront utilisées pour le développement des métriques. La Figure 2.4 donne le mode de fonctionnement de ce type de métriques.

Ye et al [111] ont construit un codebook à partir des blocs de l'image où la distribution des mots de ce dernier a été examinée. En fonction de cette distribution, des informations importantes de la qualité perçue peuvent être extraites. La métrique comprend trois étapes ; l'extraction des caractéristiques locales, la construction de dictionnaire qui représente l'image et la régression pour obtenir le score de qualité.

Le filtre de Gabor a été choisi pour l'extraction des caractéristiques. L'utilisation de ce type de filtre a été motivée par le fait que les images qui subissent la même dégradation et qui ont une qualité similaire partagent des textures identiques.

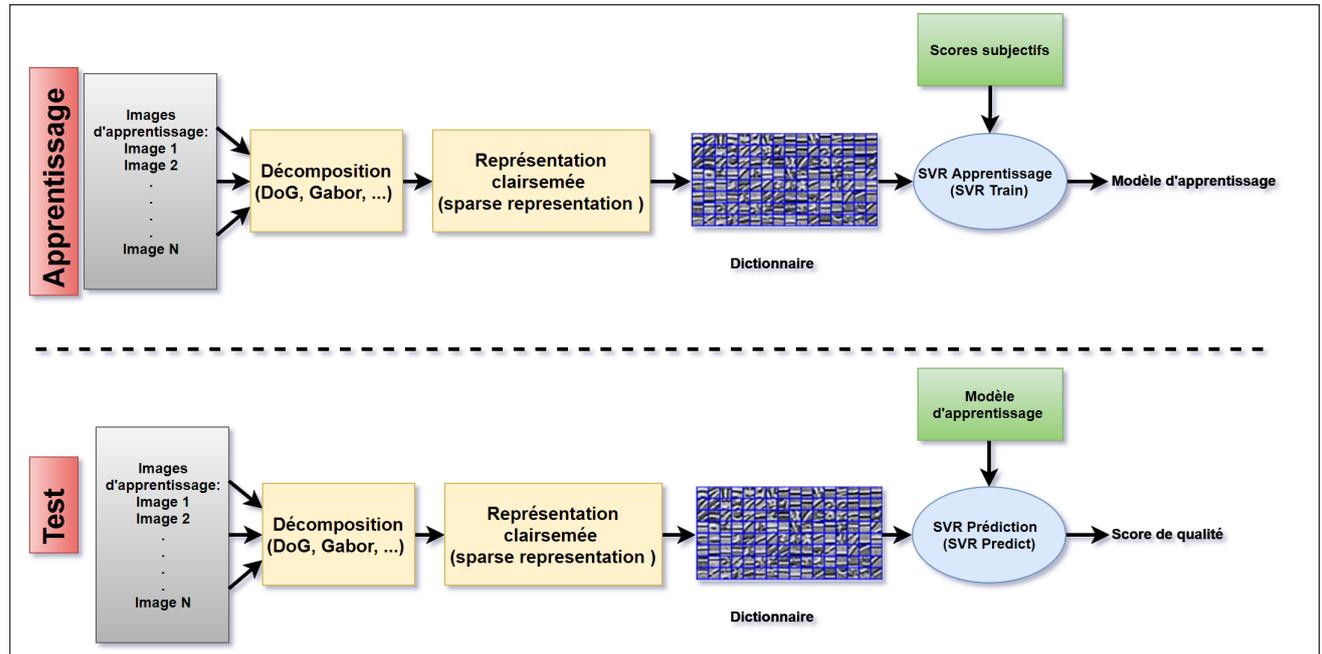


FIGURE 2.4: Métriques basées sur les dictionnaires (codebook).

L'étape suivante consiste à construire un dictionnaire à partir des blocs de l'image filtrée. Un vecteur de caractéristiques de filtre Gabor est obtenu. Le dictionnaire est ensuite créé en utilisant l'algorithme de clustering k-means.

Pour la régression, deux approches ont été adoptées, la première consiste à faire une cartographie (mapping) directe entre les mots du dictionnaire et leurs scores subjectifs. Dans la deuxième approche, les histogrammes des mots du dictionnaire ont été utilisés comme entrée pour le SVR. Une autre approche MSE-EQI utilisant les codebook a été proposée par P.Ye et al [112]. Une décomposition uniforme et aléatoire a été faite pour les images d'apprentissage. Les blocs obtenus sont ensuite normalisés. Le codebook est construit en utilisant l'algorithme de clustering k-means. Il a été normalisé de telle sorte que chaque composante aura une taille unique. Les mots du codebook ont pu montrer les différentes distorsions existantes. Ensuite, le codage d'affectation (soft-assignment coding) et la mise en commun maximale (max pooling) sont utilisés pour coder ces caractéristiques. Le score final est donné par la machine à vecteur de support (SVR) avec un noyau linéaire. Elle a été adoptée pour surmonter les non-linéarités introduites dans les caractéristiques au cours de la phase de mise en commun.

Les caractéristiques d'entropie spatiale et spectrale ont été utilisées dans la métrique [49]. Dans cette mesure l'évaluation a été faite sur deux étapes, la première consiste à classifier les types de distorsion affectant l'image. La deuxième étape consiste à mesurer la qualité selon le type de distorsion trouvé. L'entropie locale peut capturer la structure et le niveau de dépendance entre les pixels de l'image. Un sous échantillonnage de l'image est premièrement appliqué.

Ensuite, l'image a été décomposée en bloc ($M \times M$) pour calculer l'entropie locale spatiale et fréquentielle. Les entropies ont été triées dans l'ordre ascendant. Pour créer le lien entre les caractéristiques, la méthode de pooling a été utilisée. Une probabilité de classification est formée pour chaque distorsion, puis la fonction de régression est obtenue pour chaque type de distorsion en utilisant les scores subjectifs. Cela donne deux vecteurs : le vecteur de distorsion et le vecteur de distorsion de qualité. Le score final de qualité est obtenu par le calcul de produit de ces deux vecteurs.

2.4.4 Limitations des métriques sans référence existantes

Les métriques MSR-EQI représentent un avancement important dans le domaine de qualité en les comparant aux métriques MAR-EQI et MRR-EQI. Néanmoins, un long chemin et un effort considérable sont encore devant ces métriques pour qu'elles répondent parfaitement aux différents besoins. Pour cela, différentes limitations peuvent être encore enregistrées. Nous allons en présenter quelques unes dans ce paragraphe.

Les métriques basées sur NSS et celles basées sur les réseaux de neurones ont des limitations communes. Parmi les limitations majeures, il y a la dépendance aux bases d'apprentissage. Nous avons vu dans les paragraphes précédents que ces métriques sont basées sur deux étapes. Une pour l'apprentissage et la deuxième pour le test. Or, pour la première étape un nombre important d'images est utilisé pour familiariser le SVR respectivement le réseau de neurone aux différents comportements des dégradations pour l'image. D'après les résultats d'expérimentation de ces métriques, la performance se dégrade drastiquement si la mesure est appliquée sur une autre base.

Pour les métriques basées sur les codebooks, elles partagent les mêmes limitations précédentes. En plus, le grand nombre de caractéristiques utilisées pour construire les codebooks pose un grand problème. En effet, il se répercute directement sur le temps de calcul de la métrique et il augmente la complexité.

2.5 Conclusion

Dans ce chapitre, un état de l'art a été présenté sur les métriques de qualité d'évaluation d'images existantes dans la littérature. Un rappel sur les différentes métriques dans littérature a été donné.

Dans un premier temps, les métriques d'évaluation de qualité avec référence MAR-EQI ont été rapportées. Ensuite, une description succincte des métriques avec références réduites MRR-EQI a été donnée. On a pu constater que l'application de ces deux types de mesures pose souvent un problème. Le fait qu'elles soient liées à l'image de référence d'une manière ou d'une autre bloque leur utilisation pour des applications temps réel.

CHAPITRE 2. MÉTRIQUES DÉVELOPPÉES POUR L'ÉVALUATION DE LA QUALITÉ D'IMAGE : ÉTAT DE L'ART

Les MSR-EQI sont les successeurs des MAR et MRR au regard de leurs flexibilités et dépendances. Un état de l'art des travaux représentatifs a été rapporté. Nous avons mis en évidence une classification selon les techniques utilisées pour l'évaluation. De ce fait, les métriques ont été regroupées en trois classes ; les MSR basées sur les NSS, les MSR basées sur les réseaux de neurones et les MSR basées sur les dictionnaires. Au final, les limitations et les points faibles de chaque technique ont été examinées. Dans le chapitre suivant, nous allons présenter nos contributions. Celles-ci traitent des nouvelles techniques pour l'évaluation de la qualité d'image sans référence .

FUSION DES SCORES PAR RÉGRESSION EN UTILISANT LES SÉPARATEURS À VASTE MARGE (SVR)

Sommaire

3.1	Introduction	46
3.2	Machines à Vecteurs de Support (Support vector regression (SVR))	47
3.2.1	Machines à Vecteurs de Support (SVR) linéaire	47
3.2.2	Machines à Vecteurs de Support (SVR) non-linéaire	50
3.3	Évaluation de la qualité d'image par fusion de scores	51
3.3.1	Scores de qualité	51
3.3.2	Fusion des scores	53
3.4	Résultats et Discussions	54
3.5	Conclusion	58

3.1 Introduction

L'omniprésence de l'information visuelle (sous forme d'images et de vidéos) dans tous les secteurs a besoin de méthodes fiables et efficaces pour évaluer sa qualité. Par conséquent, de nombreux chercheurs se sont concentrés à développer des métriques de qualité. L'EQI peut être organisée en trois grandes classes. La première classe comporte les métriques avec référence complète (MAR-EQI). Pour cette classe, l'image de référence est utilisée avec l'image dégradée, où on cherche à quantifier la moindre différence entre les deux. Cette quantification est transformée à un indice de qualité.

La deuxième classe est l'évaluation avec référence réduite (MRR-EQI). Pour celle-ci nous n'avons pas accès à la globalité de l'image de référence, mais plutôt à une partie. Selon ce qu'on possède de la référence, les mêmes informations sont extraites de l'image dégradée. Une comparaison est effectuée pour évaluer la qualité de l'image en question.

La troisième classe est dédiée à l'évaluation sans référence (MSR-EQI). Dans ce type d'évaluation, on ne possède que l'image dégradée. Cela rend plus difficile ce genre d'évaluation mais elle reste la plus recherchée. Ceci est due au fait que dans la majorité des applications l'image de référence est inexistante.

Dans ce chapitre, nous allons proposer une métrique sans référence MSR-EQI, qui cherche à surmonter les difficultés qui existent dans différentes métriques de l'état de l'art. L'algorithme utilise cinq MSR-EQI pour calculer un vecteur de score, qui est utilisé à l'entrée de la machine à vecteur de support (SVR). Après la phase d'apprentissage, le SVR sera capable d'évaluer toute image de test.

Les machines à vecteur de support, machine learning, ont été très utilisées dans la qualité d'image à la fois pour la classification et la régression. Ceci est motivé par la performance de ces algorithmes et les résultats qu'elles donnent.

Dans [63], Narwaria et Lin ont proposé une approche basée sur les machines par apprentissage pour évaluer la qualité perspective. La métrique utilise la décomposition en valeurs singulières (SVD). Celles-ci sont utilisées comme caractéristiques pour capturer les informations structurelles dans les images. Et pour prédire le score de la qualité perceptuelle, une régression à l'aide de machines à vecteurs de supports (SVM) est utilisée pour le processus de regroupement (pooling) des caractéristiques. Les auteurs n'ont pas fourni des comparaisons avec d'autres métriques.

Les auteurs de [89], ont proposé une métrique qui utilise les machine learning pour mesurer la qualité visuelle des images codées par la compression JPEG. Les caractéristiques sont extraites en tenant compte des facteurs liés à la sensibilité visuelle humaine, tels que la longueur du contours (bord), l'amplitude du bord, la luminance du fond et l'activité du fond. La qualité visuelle d'une image est ensuite calculée en utilisant le nombre de classes prédites et leur probabilité estimée.

3.2 Machines à Vecteurs de Support (Support vector regression (SVR))

Les machines à vecteurs de support ou séparateurs à vaste marge, support vector machine (SVM) en anglais ; sont des techniques destinées à résoudre des problèmes de discrimination et de régression. Les SVM ont été développées dans les années 1990 à partir des considérations théoriques de **Vladimir Vapnik** [6] dans un contexte purement industriel. Parmi les premiers domaines d'application des SVM, on a le domaine de reconnaissance optique de caractères. Dans ce cadre les classificateurs SVM sont devenus un outil concurrentiel aux meilleurs systèmes disponibles. Aussi les SVM ont été utilisées pour les problèmes de régression, support vector regression SVR, où d'excellentes performances ont été obtenues. Par la suite, une description des machines à vaste marge pour la régression est donnée [78]. Elles seront utilisées dans la métrique proposée.

3.2.1 Machines à Vecteurs de Support (SVR) linéaire

Supposons un ensemble de données $\{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\} \in \mathbb{R}^N \times \mathbb{R}$, où x_i sont les caractéristiques et y_i sont les valeurs cibles (targets). Dans la régression ε -SVR, l'objectif est de trouver une fonction $f(x)$ qui a comme écart ε par rapport aux cibles obtenues y_i pour toutes les données d'entraînement. Autrement dit, toute erreur est permise tant qu'elle sera inférieure à ε , mais aucune erreur supérieure à celui-ci ne sera acceptée. Le SVR dans sa forme linéaire peut être exprimé comme suit :

$$(3.1) \quad f(x) = \langle w, x \rangle + b \text{ avec } w \in \mathbb{R}^N, b \in \mathbb{R}$$

où w le vecteur des poids et b une constante à déterminer. Cette fonction linéaire doit avoir une certaine platitude et satisfaire la condition suivante pour chaque valeur cible y_i [88] :

$$(3.2) \quad \|\langle w, x \rangle + b - y_i\|_2 \leq \varepsilon, \forall i = 1, \dots, N.$$

où $\|\cdot\|_2$ est la norme euclidienne.

Pour cela, on doit trouver des poids w minimas qui assurent la platitude de la fonction f . En d'autres termes, la norme $\|w\|_2$ est minimisée sachant que les erreurs doivent être inférieures à ε .

$$(3.3) \quad \begin{array}{ll} \text{Min} & \frac{1}{2} \|w\|_2^2 \\ \text{Sujet de} & \|\langle w, x \rangle + b - y_i\|_2 \leq \varepsilon, \forall i = 1, \dots, N. \end{array}$$

Cette formulation de problème considère l'existence d'une fonction linéaire f qui approxime toutes les données (x_i, y_i) avec une précision ε . Mais dans la pratique cela n'est pas toujours le

CHAPITRE 3. FUSION DES SCORES PAR RÉGRESSION EN UTILISANT LES SÉPARATEURS À VASTE MARGE (SVR)

cas. En effet, une certaine tolérance est introduite en utilisant le concept de la marge souple, soft margin en anglais.

L'introduction des variables de relâchement ξ_i, ξ_i^* transforme le problème d'optimisation comme suit :

$$\begin{aligned}
 \text{Min} \quad & \frac{1}{2} \|w\|_2^2 + C \sum_{i=1}^N (\xi_i + \xi_i^*) \\
 \text{Sujet de} \quad & y_i \leq \langle w, x \rangle + b + \varepsilon + \xi_i \\
 & y_i \geq \langle w, x \rangle + b - \varepsilon - \xi_i^* \\
 & \xi_i \geq 0, \xi_i^* \geq 0, i = 1, \dots, N.
 \end{aligned}
 \tag{3.4}$$

où C est un paramètre de pénalité définissant l'écart toléré de l'erreur ε , ce qui induit l'ajustement de la quantité d'erreur envers la platitude de la fonction f . Cela correspond au problème ε -insensible $|\xi|_\varepsilon$:

$$|\xi|_\varepsilon = \begin{cases} 0, & |\xi| \leq \varepsilon \\ |\xi| - \varepsilon, & \text{ailleurs.} \end{cases}
 \tag{3.5}$$

La Figure 3.1 montre le concept de la marge souple ε -insensible.

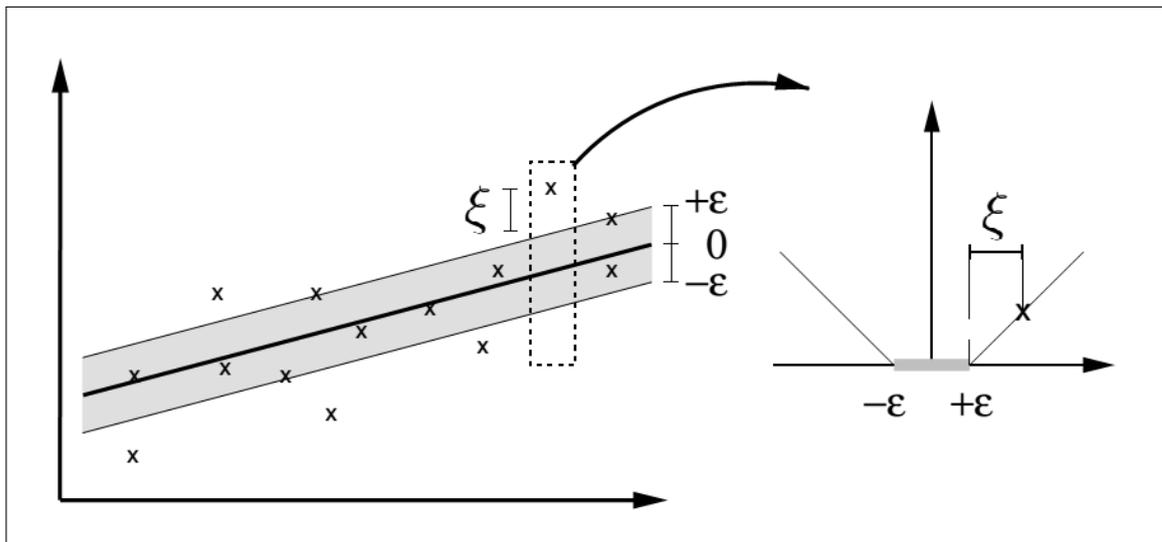


FIGURE 3.1: Concept du marge souple (soft margin) dans le cas linéaire.

Le problème peut être résolu en utilisant le Lagrangien avec les contraintes correspondantes en introduisant un groupe de variables duales :

$$\begin{aligned}
 L &:= \frac{1}{2} \|w\|_2^2 + C \sum_{i=1}^N (\xi_i + \xi_i^*) - \sum_{i=1}^N (\eta_i \xi_i + \eta_i^* \xi_i^*) \\
 &- \sum_{i=1}^N \beta_i (\varepsilon - \xi_i - y_i + \langle w, x \rangle + b) \\
 &- \sum_{i=1}^N \beta_i^* (\varepsilon - \xi_i^* + y_i - \langle w, x \rangle + b)
 \end{aligned}$$

(3.6)

avec $(\eta_i, \eta_i^*, \beta_i, \beta_i^*)$ les multiplicateurs de Lagrange. Ces variables duales doivent satisfaire la contrainte de positivité. Les dérivées partielles par rapport à (w, b, ξ_i, ξ_i^*) sont données comme suit :

$$(3.7) \quad \partial_b L = \sum_{i=1}^N (\beta_i^* - \beta_i) = 0$$

$$(3.8) \quad \partial_w L = w - \sum_{i=1}^N (\beta_i^* - \beta_i) x_i = 0$$

$$(3.9) \quad \partial_{\xi_i} L = C - \beta_i - \eta_i = 0$$

$$(3.10) \quad \partial_{\xi_i^*} L = C - \beta_i^* - \eta_i^* = 0$$

En substituant les équations (3.7), (3.8), (3.9) et (3.10) dans l'équation (3.6) on obtient le problème d'optimisation suivant :

$$\begin{aligned}
 (3.11) \quad \text{Max} &- \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N (\beta_i - \beta_i^*)(\beta_j - \beta_j^*) \langle x_i, x_j \rangle \\
 &- \varepsilon \sum_{i=1}^N (\beta_i + \beta_i^*) + \sum_{i=1}^N (\beta_i - \beta_i^*) y_i \\
 \text{Sujet de} &\sum_{i=1}^N (\beta_i - \beta_i^*) = 0 \text{ et } \beta_i, \beta_i^* \in [0, C]
 \end{aligned}$$

En dérivant l'équation (3.6), les poids peuvent être déterminés à partir de l'équation (3.8) :

$$(3.12) \quad w = \sum_{i=1}^N (\beta_i - \beta_i^*) x_i$$

et la fonction f devient :

$$(3.13) \quad f(x) = \sum_{i=1}^N (\beta_i - \beta_i^*) \langle x_i, x \rangle + b$$

Le paramètre de biais b est calculé en utilisant les conditions de Karush-Kuhn-Tucker (KKT), pour lesquelles le produit entre les variables duales et les contraintes doit être nul :

$$\begin{aligned}
 \beta_i(\varepsilon + \xi_i + f(x_i) - y_i) &= 0 \\
 \beta_i^*(\varepsilon + \xi_i^* + f(x_i) + y_i) &= 0 \\
 (C - \beta_i)\xi_i &= 0 \\
 (C - \beta_i^*)\xi_i^* &= 0
 \end{aligned}
 \tag{3.14}$$

On peut calculer b comme suit :

$$b = \begin{cases} y_i - \varepsilon - wx_i, & \beta_i < C \\ y_i + \varepsilon - wx_i, & \beta_i^* < C \end{cases}
 \tag{3.15}$$

Les vecteurs de support sont définis comme les points qui contribuent aux prédictions données par (3.13) et les x_i où $\beta_i - \beta_i^* \neq 0$. Ces x_i représentent les Vecteurs de Support.

3.2.2 Machines à Vecteurs de Support (SVR) non-linéaire

Pour le cas non linéaire, on doit suivre les mêmes étapes que dans le cas linéaire. Néanmoins, un changement d'espace pour les données x_i est nécessaire. Cela est dû à la projection des données dans un espace de dimension supérieur qui rend le problème complètement linéaire. La projection est faite comme suit $\varphi : \mathbb{R}^N \rightarrow F$, où F est appelé espace de données. Après projection, le produit scalaire dans (3.1) devient $k(x_i, x) = \langle \varphi(x_i), \varphi(x) \rangle$, qui est le produit scalaire entre les images des données dans le nouvel espace F . Les poids sont donnés par :

$$w = \sum_{i=1}^N (\beta_i - \beta_i^*) \varphi(x_i)
 \tag{3.16}$$

et l'équation (3.13) est donnée à nouveau :

$$f(x) = \sum_{i=1}^N (\beta_i - \beta_i^*) k(x_i, x) + b
 \tag{3.17}$$

La fonction $k(x_i, x)$ est appelée fonction noyau. Les noyaux les plus utilisés sont définis comme suit :

- **Noyau Linéaire** : $k(x_i, x_j) = \langle x_i, x_j \rangle$;
 - **Noyau Polynomial** : $k(x_i, x_j) = (\gamma \langle x_i, x_j \rangle + r)^d, \gamma > 0$;
 - **Noyau à base Radial (RBF)** : $k(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2), \gamma > 0$;
 - **Noyau Sigmoidé** : $k(x_i, x_j) = \tanh(\gamma \langle x_i, x_j \rangle + r)$;
- où γ, r et d sont les paramètres du noyau.

3.3 Évaluation de la qualité d'image par fusion de scores

Malgré la multitude des métriques d'évaluation de qualité proposées, il n'existe pas une métrique universelle qui puisse fonctionner correctement pour toute type de distorsion. À partir de là, développer une MSR-EQI capable d'évaluer différents types de distorsion est très demandé. La méthode proposée est conçue sur le principe de fusion. Elle fonctionne de la manière suivante, pour chaque image, différents scores de qualité ont été obtenus à partir des métriques existantes. Par la suite, une fusion de ces scores est effectuée. On a utilisé les machines à vecteur de support SVR pour obtenir le score final. La Figure 3.2 représente la cartographie de la métrique.

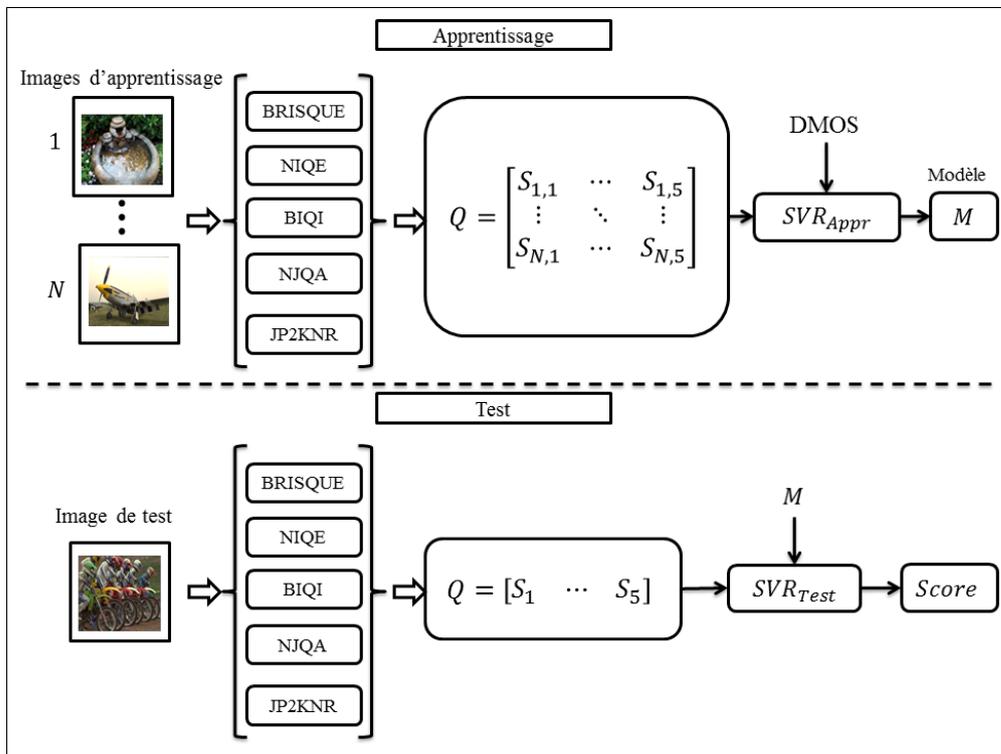


FIGURE 3.2: Diagramme de la métrique proposée

3.3.1 Scores de qualité

Pour obtenir des scores de qualité, cinq métriques ont été utilisées en se basant sur leurs performances.

- **Métrique 1 :** Blind/Reference-less Image Spatial Quality Evaluator (BRISQUE) [56] : L'approche peut être résumée comme suit. À partir d'une image dégradée, les luminances normalisées localement sont calculées. Ces luminances ont des propriétés statistiques qui peuvent être modifiées par la présence de distorsion. Une modélisation par une loi

Gaussienne a été adoptée, et toute déviation de ce modèle est utilisée pour la quantification de ces changements afin de permettre la prédiction de la qualité perceptuelle.

- **Métrique 2** : Natural Image Quality Evaluator (NIQE) [57] : La méthode est basée sur la construction d'une collection de caractéristiques de qualité puis l'adaptation de celle-ci à un modèle Gaussien multivarié. Les caractéristiques de qualité sont dérivées à partir d'un modèle simple mais très régulier de statistiques de scènes naturelles. La qualité d'une image de test est alors exprimée comme la distance entre le modèle Gaussien multivarié de cette image et un modèle généralisé construit à partir d'un corpus d'images naturelles.
- **Métrique 3** : Blind Image Quality Index (BIQI) [59] : Cette méthode évalue la qualité d'une image aveuglement sans aucune connaissance de la source de distorsion. Ceci est obtenu en utilisant des statistiques distordues de l'image, une extension des statistiques de scène naturelle pour des images déformées. Chaque dégradation a une signature unique caractérisée par les statistiques distordues. Cette signature est utilisée pour classer les images dans des catégories de distorsion et aussi pour montrer le comportement des dégradations. Dans la phase finale, une combinaison de ces statistiques est utilisée pour produire l'indice de qualité.
- **Métrique 4** : No-reference Quality Assessment Algorithm for JPEG Compressed Images (NJQA) [26] : Pour cette métrique les auteurs proposent deux étapes pour évaluer la qualité des images compressées par la compression JPEG. La première étape génère la carte de pertinence de l'image. Dans la deuxième étape, la qualité est estimée en comptant pour chaque bloc le nombre de coefficients nuls de la transformée en cosinus discrète DCT, puis en additionnant et en pondérant ces comptes sur la base de la carte de pertinence.
- **Métrique 5** : No-Reference Quality Assessment of JPEG2000 Compressed Images (JP2KNR) [86] : Cette mesure est basée sur les statistiques de scènes naturelles NSS pour mesurer la qualité des images. L'utilisation des statistiques de scènes naturelles repose sur le fait que la présence d'une dégradation perturbe ces statistiques. Cette idée est utilisée pour évaluer la qualité des images déformées par la compression JPEG2000. Où elle perturbe les dépendances non linéaires qui sont présentes dans les images naturelles. Un modèle statistique non linéaire est adopté, pour lequel toute déviation de l'allure des images naturelles est utilisée pour prédire le score final.

Dans le Tableau 3.1, on présente les performances des différentes métriques décrites précédemment. Les coefficients de corrélation linéaire PLCC pour chaque dégradation sont obtenus à partir des expérimentations sur la base d'image LIVE [87].

Tableau 3.1: Coefficients de corrélation linéaire PLCC

Métrique \ Dist	JP2K	JPEG	WN	Gblur	FF
BRISQUE	0.9229	0.9734	0.9851	0.9506	0.9030
nique	0,9370	0,9564	0.9773	0.9525	0.9128
BIQI	0.8086	0.9011	0.9538	0.8293	0.7328
NJQA	-	0,970	-	-	-
jp2knr	0.93	-	-	-	-

3.3.2 Fusion des scores

Comme il a été mentionné dans [1], les vecteurs à vaste marge SVR font partie des méthodes de fusion. Ainsi, l'utilisation du SVR pour fusionner les scores obtenus à partir des différentes mesures d'EQI est le principal objectif de cette métrique .

En appliquant les 5 métriques pour chaque image, 5 différents scores ont été obtenus. Ensuite, les images sont représentées par leurs vecteur de score Q . Où $Q = (S_1, \dots, S_M)$, avec M qui est le nombre de métriques.

Pour faire la liaison entre les machine à vaste marge SVR et notre métrique, l'ensemble de données $\{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\} \in \mathbb{R}^N \times \mathbb{R}$ devient $\{(Q_1, DMOS_1), (Q_2, DMOS_2), \dots, (Q_m, DMOS_m)\} \in \mathbb{R}^N \times \mathbb{R}$. Avec Q représente le vecteur de score et $DMOS$ représente le score subjectif associé à chaque image (Difference Mean Opinion Score DMOS). La fonction f dans le cas non linéaire est donnée par :

$$(3.18) \quad f(Q) = \sum_{i=1}^N (\beta_i - \beta_i^*) k(Q_i, Q) + b$$

Dans le cadre d'utilisation du SVR, deux étapes sont nécessaires :

- 1- **Apprentissage du SVR** : La phase d'apprentissage est cruciale pour les machines à vecteur de support. Cette étape permet d'entraîner la machine à reconnaître comment chaque distorsion affecte l'image et le vecteur de score associé à cette dégradation. À la fin, un modèle qui représente ce que peut trouver le SVR dans la phase de test est obtenu. Dans cette étape, on a besoin du vecteur de score objectif et du score subjectif qui lui correspond. L'équation (3.19) résume la procédure d'apprentissage :

$$(3.19) \quad M = SVR_{Appr}(Q, DMOS)$$

avec M le modèle d'apprentissage.

- 2- **Test en utilisant le modèle SVR** : La deuxième étape consiste à évaluer des images qui ne font pas partie du groupe d'apprentissage. Cela dit que le SVR ne connaît pas le type de dégradation affectant l'image. Comme pour la première étape, on a besoin du vecteur

de score Q . Toutefois, on ne va pas utiliser les scores subjectifs mais le modèle M obtenu précédemment :

$$(3.20) \quad \text{Score} = SVR_{Test}(Q, M)$$

avec $Score$ l'indice final qui décrit la qualité de l'image après la fusion.

3.4 Résultats et Discussions

Afin d'étudier le fonctionnement de cette nouvelle métrique, on a utilisé la base d'images LIVE [87] pour mener nos tests. La base LIVE est la base la plus utilisée pour l'évaluation de la qualité d'image.

La même procédure utilisée dans [56] est suivie, où, la base est décomposée aléatoirement en deux groupes, 80% des images sont utilisées pour l'apprentissage et 20% pour les tests. Ceci est réalisé pour s'assurer que les résultats rapportés ne dépendent pas de l'espace d'apprentissage. La procédure aléatoire d'apprentissage-test est répétée 1000 fois où la moyenne de ces itérations est donnée.

Les indices de performance utilisés sont la corrélation linéaire de Pearson (PLCC), la corrélation de Spearman (SROCC) et la moyenne quadratique de l'erreur (RMSE).

N.B avant de calculer la corrélation linéaire PLCC les scores objectifs sont passés dans une fonction logistique (logistic) non linéaire [87] comme suit :

$$(3.21) \quad \text{Quality}(x) = \beta_1 \text{logistic}(\beta_2, (x - \beta_3)) + \beta_4 x + \beta_5$$

$$(3.22) \quad \text{logistic}(\tau, x) = \frac{1}{2} + \frac{1}{(1 + \exp(\tau x))}$$

avec x score objectif, $\beta_1, \beta_2, \beta_3, \beta_4$ et β_5 sont des paramètres à définir en minimisant l'erreur. Une valeur proche de 1 pour les coefficients de corrélations indique une bonne performance de la métrique et vice-versa. Pour l'erreur, il faut que les valeurs soient minimales. Les résultats obtenus ainsi que des comparaisons avec les cinq métriques utilisées sont donnés dans les Tableaux 3.2, 3.3 et 3.4.

Tableau 3.2: Moyenne de corrélation linéaire PLCC pour 1000 itérations

Métrique \ Dist	JP2k	JPEG	WN	Gblur	FF
BRISQUE	0.9229	0.9734	0.9851	0.9506	0.9030
niqe	0,9370	0,9564	0.9773	0.9525	0.9128
BIQI	0.8086	0.9011	0.9538	0.8293	0.7328
NJQA	-	0,970	-	-	-
jp2knr	0.92	-	-	-	-
Métrique proposée	0.9673	0.9897	0.9995	0.9819	0.9414

Tableau 3.3: Moyenne de corrélation de Spearman SROCC pour 1000 itérations

Métrique \ Dist	JP2k	JPEG	WN	Gblur	FF
BRISQUE	0.9139	0.9647	0.9511	0.8768	0.9395
NIQE	0,9172	0,9382	0.9662	0.9341	0.8594
BIQI	0.7995	0.8914	0.9510	0.8463	0.7067
NJQA	-	0,967	-	-	-
JP2KNR	0.94	-	-	-	-
Métrique proposée	0.9518	0.9793	0.9908	0.9765	0.9123

Tableau 3.4: Moyenne de l'erreur quadratique RMSE pour 1000 itérations

Metric \ Dist	JP2k	JPEG	WN	Gblur	FF
BRISQUE	8.484	8.167	4.331	6.635	11.983
niqe	8.123	6.341	5.321	8.916	15.923
BIQI	14.842	13.755	8.409	10.234	19.291
NJQA	-	8.645	-	-	-
jp2knr	8.05	-	-	-	-
Métrique proposée	6,385	4,440	2,827	3,555	9,735

Les Tableaux 3.2 et 3.3 montrent que la métrique proposée donne des coefficients de corrélation plus élevés par rapport aux méthodes de l'état de l'art. En effet, on a pu voir des coefficients qui dépassent 0.94 pour la corrélation linéaire et 0.95 pour la corrélation de Spearman. On peut dire que la fusion de score est une stratégie puissante qui permet de surmonter le problème de distorsions multiples.

Les résultats de l'erreur RMSE sont donnés dans le Tableau 3.4. Encore une fois, les résultats de la métrique proposée montrent une efficacité remarquable, ils donnent les erreurs les moins élevées. Cela veut dire que les scores objectifs obtenus après fusion sont plus proches des scores subjectifs et cela implique une évaluation crédible.

Un plot qui représente l'allure des scores objectifs (*Fscores*) vis-à-vis des scores subjectifs (*DMOS*) pour toutes les dégradations est donné dans la Figure 3.3. On peut constater que les scores objectifs en bleu sont plus proches de la courbe rouge qui représente les scores subjectifs.

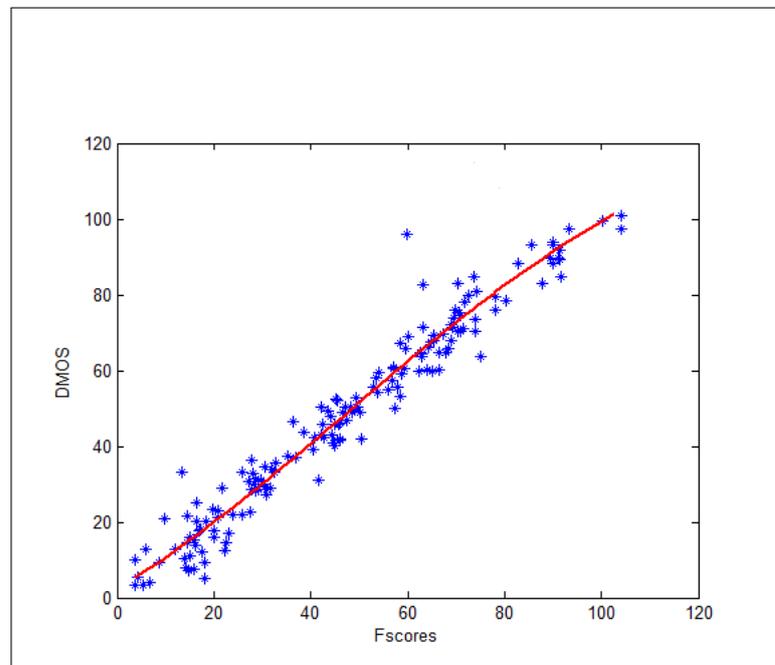


FIGURE 3.3: Graphe des scores objectifs contre les scores subjectifs

Les plots dans la Figure 3.4 représentent le nuage de point pour chaque dégradation. Elles montrent que la métrique se comporte bien, où les scores produits reflètent la performance de la nouvelle méthode.

3.4. RÉSULTATS ET DISCUSSIONS

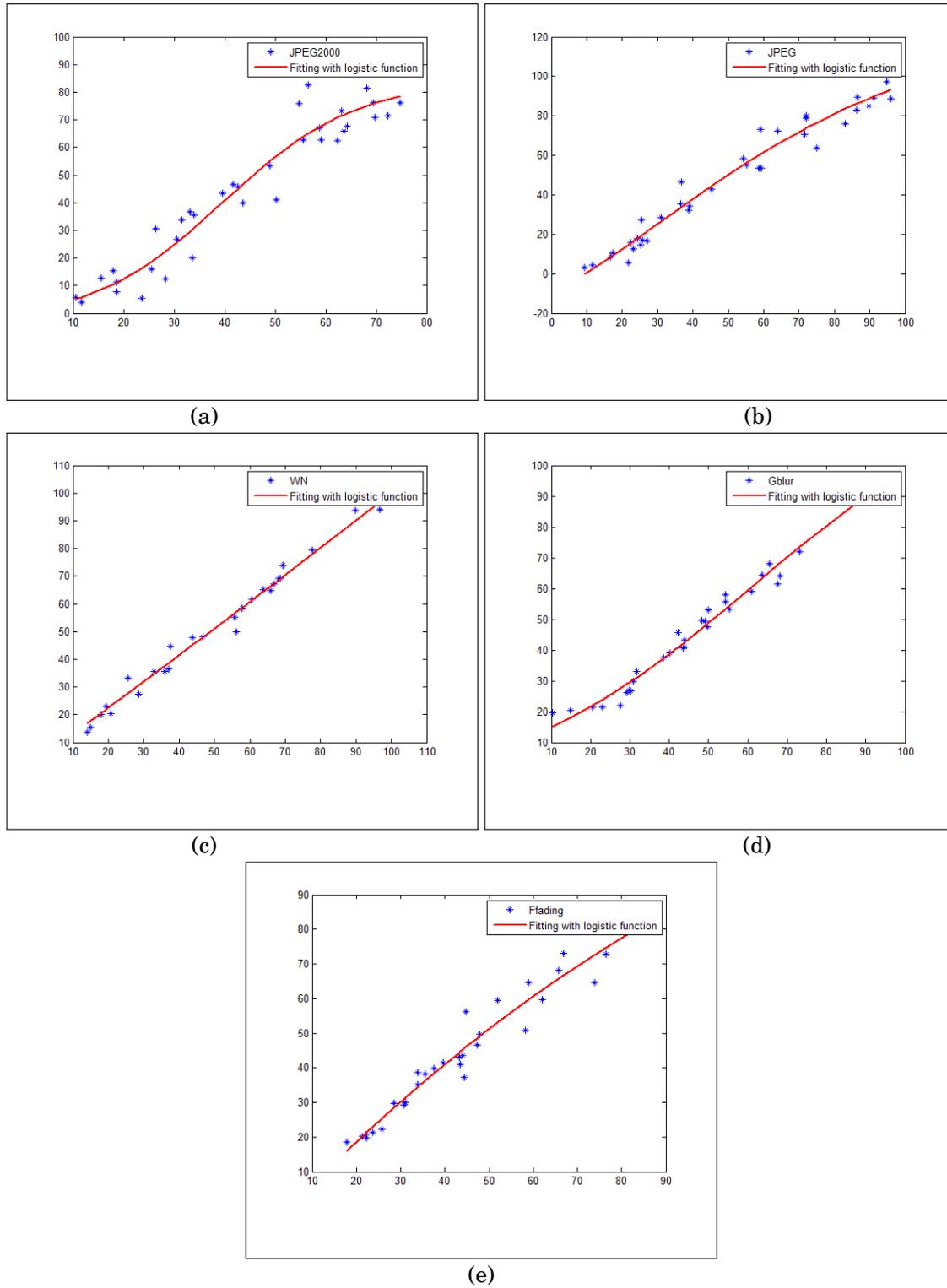


FIGURE 3.4: Plots des scores objectifs contre les scores subjectifs pour chaque dégradation, (a)JPEG2000, (b)JPEG, (c)WN, (d)Gblur, (e) FF

3.5 Conclusion

Dans ce chapitre une nouvelle métrique a été proposée. Elle est basée sur la fusion de différents scores de qualité. Cela est fait pour répondre au problème de distorsions multiples où la majorité des métriques proposées dans la littérature ont échoué. Cette nouvelle métrique utilise cinq différents scores obtenus à partir des métriques de qualité. Comme outil de fusion, les puissantes machines learning SVR ont été adoptées. La mesure proposée a donné des résultats prometteurs en termes de corrélation avec la perception humaine, et elle est très compétitive aux métriques d'EQI de l'état de l'art. Néanmoins, la métrique reste encore incomplète. Comme décrit dans les sections précédentes la fusion est faite pour toutes les métriques. Fusionner les métriques sans différences peut poser un problème majeur dans le cas où la métrique à fusionner ne correspond pas un tel type de distorsion. Cela dit, on doit avoir un algorithme de sélection qui doit choisir les métriques les plus appropriées pour chaque type de dégradation. Ainsi, cette sélection va nous permettre d'augmenter le nombre de métriques en assurant toujours la robustesse de la métrique. Cela va être abordé dans le chapitre suivant.

ÉVALUATION GÉNÉRIQUE DE LA QUALITÉ BASÉE SUR LES ESSAIS DE PARTICULES ET LES MACHINE À VECTEUR DE SUPPORT

Sommaire

4.1	Introduction	60
4.2	Aperçu sur les métriques d'évaluation basées sur la fusion	60
4.3	Méthode proposée	62
4.3.1	Étape de sélection des métriques	62
4.3.2	Étape de fusion	66
4.4	Résultats et Discussions	70
4.4.1	Indépendance de la base d'apprentissage	77
4.4.2	Comparaison avec les algorithmes génétiques	78
4.5	Conclusion	80

4.1 Introduction

Malgré, la performance croissante des technologies d'acquisition d'images et le développement continu de nouvelles méthodes et systèmes embarqués pour l'amélioration des images et des vidéos, la quantification de la qualité de l'image reste un problème difficile. En effet, les images numériques peuvent souffrir de diverses distorsions et artefacts lors de l'acquisition, de traitement, de la compression et de la transmission. Certaines métriques de qualité d'image intéressantes dédiées à des distorsions spécifiques ont été proposées. Cependant, la quantification de la qualité de l'image sans référence est toujours un problème difficile surtout lorsque différentes distorsions affectent l'image observée. L'évaluation subjective de la qualité de l'image reste le moyen le plus fiable d'évaluer la qualité de l'image. Toutefois, la réalisation d'une évaluation subjective prend du temps et consomme beaucoup de ressources. Pour ces raisons, de nombreux chercheurs se sont concentrés sur le développement de mesures objectives efficaces, compatibles avec le jugement humain.

Une métrique pour l'évaluation des images codées par JPEG a été proposée dans [89]. Les caractéristiques de la sensibilité visuelle humaine sont extraites (longueur et amplitude du bord ainsi que la luminance du fond et l'activité du signal). La qualité perceptuelle est ensuite calculée en utilisant le nombre de classes prédites et leur probabilité postérieure estimée. La machine d'apprentissage extrême (extreme learning machine ELM) est ensuite utilisée pour calculer le score final. Cette métrique est une mesure avec spécification de distorsion ; elle est dédiée uniquement aux dégradations dues à la compression JPEG.

Dans [32], les auteurs ont proposé une métrique d'évaluation de qualité d'images sans référence MSR-EQI basée sur l'importance des parties saillantes dans l'image. Les régions importantes sont trouvées à travers une carte de saillance. Les blocs de l'image sont alors extraits de ces parties saillantes. Dans la seconde étape, l'analyse par composante zéro (ZCA) est utilisée pour normaliser les blocs sélectionnés. Pour obtenir l'indice de qualité final, un codage d'affectation (soft assignment) est utilisé pour construire un vecteur caractéristique, qui est utilisé comme entrée pour le régression du vecteur de support (SVR). La métrique n'est testée que sur la base d'image LIVE et aucune comparaison avec d'autres métriques n'a été faite.

Plusieurs mesures MSR-EQI sont conçues à partir de différents points de vues (spécification et non spécification de distorsion) et sont fusionnées d'une manière appropriée afin d'obtenir une métrique de qualité d'image qui peut couvrir différents types de dégradation. Ces stratégies visent à compenser les limites d'une seule métrique pour certaines distorsions. Dans le paragraphe suivant, un état de l'art des métriques basées sur la fusion est donné.

4.2 Aperçu sur les métriques d'évaluation basées sur la fusion

Plusieurs métriques EQI, avec ou sans référence, fondées sur le principe de la fusion ont été proposées dans la littérature. Il est reconnu qu'une seule métrique ne permet pas de quantifier

efficacement les différentes distorsions et artefacts de l'image. Pour résoudre le problème de multiples distorsions, des méthodes d'évaluation intéressantes, basées sur des schémas de fusion ont été proposées.

Dans [50], une métrique MAR-EQI basée sur la fusion supervisée a été proposée. La métrique est basée sur l'analyse de corrélation canonique (ACC). Tout d'abord, un vecteur contenant certaines métriques sélectionnées et leurs scores subjectifs correspondants est formé. Dans la deuxième, les vecteurs canoniques projectifs ainsi que leurs caractéristiques de corrélation canonique sont obtenus. L'indice de qualité final est estimé à l'aide des coefficients obtenus par le ACC et les paramètres de régression de l'erreur des moindres carrés. Cette métrique a donné de bons résultats. Dans [36], une approche basée sur la fusion multi-métrique a été proposée. La métrique détecte automatiquement le contenu de chaque bloc et le type de distorsion présent en utilisant l'apprentissage automatique (machine learning). Ensuite, les blocs sont classés en trois types ; bloc lisse, bloc de bord (contour) et bloc de texture. Dans l'étape suivante, une métrique de qualité d'image appropriée est adoptée pour chaque bloc en considérant son contenu et les types de distorsion présentent. Dans la dernière étape, tous les score des métriques de qualités appliquées sur ces blocs sont fusionnés pour obtenir le score final de qualité. Cette approche a été testée uniquement sur la base d'image TID, les résultats semblent être efficaces par rapport à l'état de l'art.

Une autre approche basée sur la classification des distorsions et l'apprentissage automatique a été proposée dans [13]. L'idée de base est de considérer les métriques de qualité comme des caractéristiques et les distorsions comme des classes. Le problème d'EQI est alors exprimé comme un problème d'identification de distorsion en utilisant un processus de classification. La classification des distorsions est effectuée en utilisant l'analyse linéaire discriminante (ALD). Et pour chaque type de dégradation, les métriques les plus appropriées sont sélectionnées pour quantifier la qualité perceptuelle. Il a été prouvé que cette méthode pourrait évaluer diverses distorsions. Des résultats efficaces sont obtenus pour la détection et l'évaluation.

Liu et al [51] ont proposé une métrique EQI avec référence basée sur la fusion multi-méthode combinée à une approche d'apprentissage (machine learning). Le score global est obtenu en combinant différentes méthodes avec des poids appropriés en utilisant un SVR. Les images déformées sont regroupées en trois à cinq groupes en fonction du type de distorsion. Une régression est ensuite effectuée pour chaque groupe. Pour minimiser la complexité, un algorithme de sélection est utilisé pour réduire les méthodes candidates. L'algorithme a fourni de bons résultats sur différentes bases d'images.

Une récente métrique basée sur la fusion a été proposée dans [65]. L'évaluation de la qualité perceptuelle de l'image est exprimée comme un problème d'optimisation. Où la principale contrainte était comment choisir les meilleurs métriques à fusionner. L'auteur a adopté les algorithmes génétiques comme solution avec des fonctions objectives bien précises. Les résultats rapportés ont montré que cette métrique dépasse plusieurs métriques dans la littérature. Outre ces métriques MAR-EQI fondées sur le principe de fusion et d'apprentissage, d'autres métriques sans référence

ont été proposées.

Dans [15] Ciancio et al. ont proposé une métrique MSR-EQI basée sur la fusion. Ils ont utilisé certaines métriques de flou et quelques caractéristiques de l'image afin de les fusionner en utilisant les réseaux de neurones. Les poids synaptiques du réseau neuronal sont obtenus sur la base d'un ensemble d'apprentissage. Une fois l'étape d'entraînement réalisée, le réseau devrait mapper ou accorder les données d'entrées à un score final de qualité. La métrique a montré des valeurs de corrélation élevées comparées à d'autres mesures de flou. Cependant, cette méthode est limitée à la distorsion de flou.

4.3 Méthode proposée

La méthode proposée utilise deux étapes pour évaluer la qualité perceptuelle de l'image. Premièrement, les métriques d'EQI les plus appropriées sont sélectionnées en utilisant l'algorithme d'optimisation par essaim particulaire (PSO) [19]. Ensuite, les mesures sélectionnées sont fusionnées en utilisant la machine learning SVR pour obtenir l'indice de qualité final. l'organigramme de la Figure 4.1 montre la métrique proposée.

4.3.1 Étape de sélection des métriques

La première étape de la méthode proposée est consacrée à la sélection des métriques à fusionner. Pour cela, nous avons adopté l'algorithme d'optimisation par essaim particulaire, *particle swarm optimization* (PSO), pour résoudre le problème de sélection des métriques.

Il a été inspiré du comportement social des animaux qui évoluent en essaim, comme les bancs de poissons et les vols groupés d'oiseaux. Kennedy et Eberhart [19] ont pu observer une dynamique de déplacement chez ces animaux un peu particulière. Lors de son déplacement, l'individu peut maintenir la cohésion de l'essaim à travers des règles très simples. Comme par exemple : rester plus proche des autres individus, aller dans la même direction que le groupe ou bien voler avec la même vitesse. Toute cette harmonie est assurée avec une simple information locale d'une part et d'autre part avec la mémoire de chaque individu. Cette information est utilisée pour prendre la décision du déplacement.

Chaque particule possède une position et une vitesse. En plus de ça, la mémoire de chacune de ces particules lui permet de se souvenir de sa meilleure position et de celle des particules voisines.

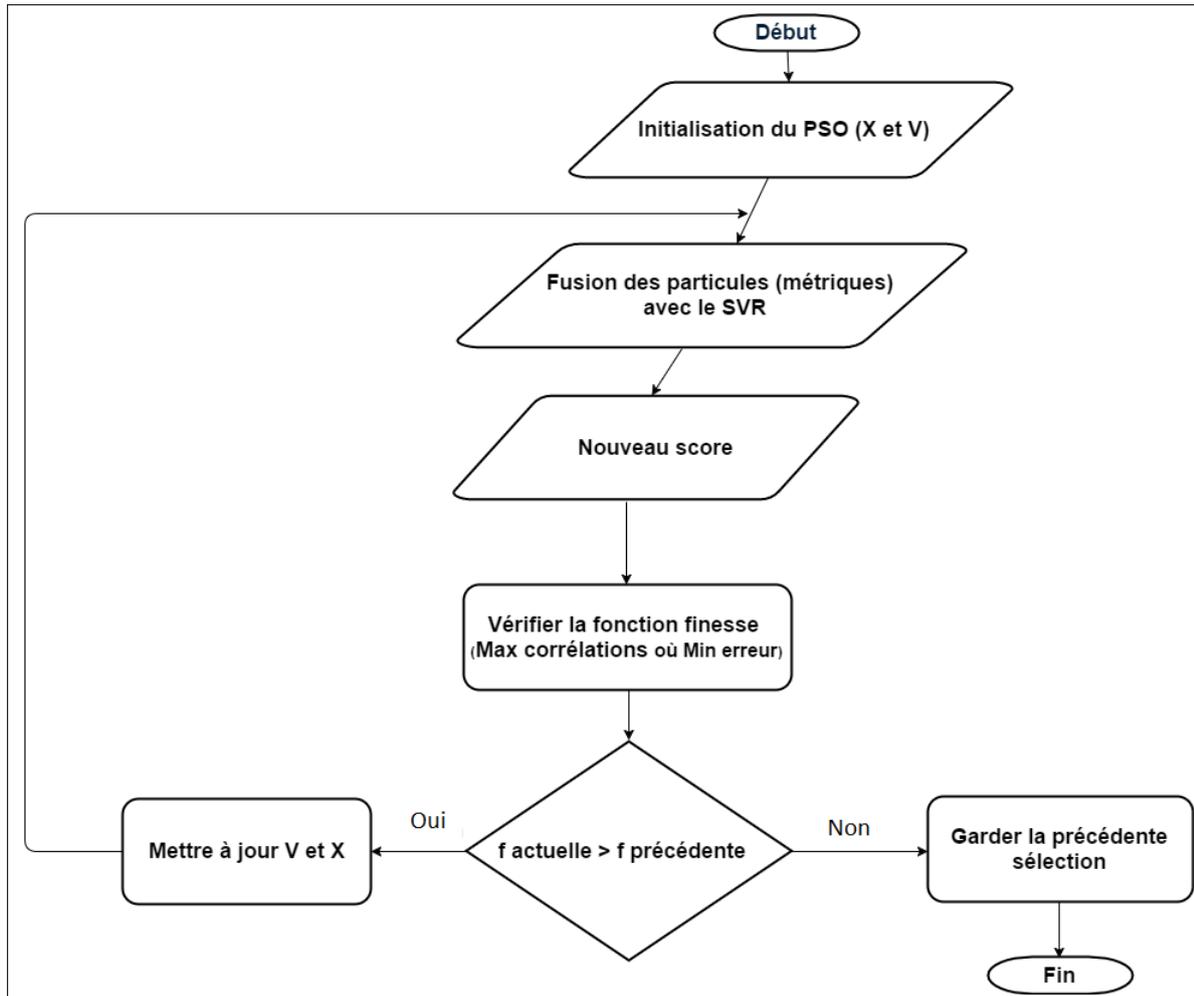


FIGURE 4.1: synoptique de la métrique proposée.

Par la suite, une description du PSO utilisé dans la conception de notre métrique de qualité d'image est donnée.

Soit $X_i = (x_1, x_2, \dots, x_M)$ un vecteur de M éléments qui indique la position de chaque particule pour notre cas métrique EQI ; où X_i est un vecteur de zéro et uns ; un veut dire que la métrique est sélectionnée pour qu'elle soit fusionnée. Chaque déplacement de particule est contrôlé avec une certaine vitesse $V_i = (v_1, v_2, \dots, v_M)$. A chaque itération k , les métriques sont fusionnées en fonction de la position des particules X_i . Les particules gardent en mémoire la meilleure position par laquelle elles sont déjà passées c.à.d la meilleure position locale $LP_i = (lp_1, lp_2, \dots, lp_M)$. La meilleure position atteinte par toutes les particules de l'essaim est donnée par $GP_i = (gp_1, gp_2, \dots, gp_M)$.

La vitesse et la position sont mises à jour après chaque itération en utilisant les équations suivantes :

$$(4.1) \quad V_i^{k+1} = \omega V_i^k + r_1 \alpha (LP_i - X_i^k) + r_2 \beta (GP_i - X_i^k)$$

$$(4.2) \quad X_i^{k+1} = X_i^k + V_i^{k+1}$$

$i = 1, 2, \dots, M$; où M indique le nombre de particules (métriques) et ω est le coefficient d'inertie. Les deux paramètres α et β sont respectivement les paramètres cognitifs et sociaux, et (r_1, r_2) sont deux nombres aléatoires répartis uniformément dans $[0, 1]$.

Le paramètre ω contrôle l'influence de la direction sur le déplacement futur.

Le paramètre α contrôle le comportement cognitif de la particule.

Le paramètre β contrôle l'aptitude sociale de la particule.

Pour chaque itération (k) (déplacement), les nouvelles positions sont évaluées à travers la fonction objective (finesse) et les deux positions LP_i et GP_i sont mis à jour comme suit :

$$(4.3) \quad LP_i(k+1) = \begin{cases} f(k+1), & \text{si } f(k+1) \geq LP_i(k) \\ LP_i(k), & \text{sinon} \end{cases}$$

$$(4.4) \quad GP_i(k+1) = \arg \max LP_i(k+1)$$

Appliquer l'algorithme PSO dans sa forme continue n'est pas très intéressant pour notre métrique. Il existe une autre forme bien adaptée à notre problème, c'est la forme binaire [29]. Cette forme est la même que la première, la seule différence est la représentation des particules qui prennent la forme de 0 ou 1. Ainsi la mise à jour de la vitesse et la position est effectuée en utilisant les deux équations suivantes :

$$(4.5) \quad V_{ri}^{k+1} = \frac{1}{1 + e^{-V^{k+1}}}$$

$$(4.6) \quad X_i^{k+1} = \begin{cases} 1, & \text{if } r < V_{ri}^{k+1} \\ 0, & \text{otherwise.} \end{cases}$$

où r est un nombre aléatoire réparti uniformément dans $[0, 1]$. Cette forme va assurer la bonne sélection des métriques. De plus, il existe un autre élément à prendre en considération : la fonction objective (finesse) qui va être responsable de la sélection des métriques.

4.3.1.1 Fonction objective (finesse)

Pour mesurer la performance d'une métrique de qualité d'image, on utilise souvent des indices qui nous permettent de juger la robustesse de cette dernière. Ces indices sont calculés en utilisant les scores objectifs et subjectifs. Parmi ces indices, on peut citer les corrélations et l'erreur. Afin

de bien les exploiter dans notre métrique, ils sont utilisés comme fonction objective. Le but est la maximisation quand il s'agit d'une corrélation et la minimisation quand il s'agit d'une erreur. Trois indices de performance ont été utilisés, on les décrit comme suit :

- la corrélation de Pearson (PLCC) : ce coefficient de corrélation permet d'analyser les relations linéaires entre les scores objectifs et subjectifs.
- la corrélation de Spearman (SROCC) : ce coefficient permet d'étudier la conformité entre les ordres de rang des scores prédictifs et subjectifs.
- Erreur quadratique moyenne (RMSE) : ce coefficient mesure l'écart absolu entre les scores objectifs et les scores subjectifs.

Chacun de ces coefficients est introduit comme fonction objectif f dans l'algorithme PSO, afin d'assurer une meilleure sélection du vecteur (X) (métrique). Pour cela, on définit trois fonctions objectives comme suit :

- **La fonction objective PLCC**

$$(4.7) \quad f_1 = PLCC(FS, D)$$

$$PLCC = \frac{\sum_{i=1}^N (FS_i - \bar{FS})(D_i - \bar{D})}{\sqrt{(\sum_{i=1}^N (FS_i - \bar{FS})^2)(\sum_{i=1}^N (D_i - \bar{D})^2)}}$$

où N est le nombre d'image utilisé. FS, D, \bar{FS} et \bar{D} sont le score objectif après fusion, le score subjectif et leurs valeurs moyennes respectivement.

- **La fonction objective SROCC**

$$(4.8) \quad f_2 = SROCC(FS, D)$$

$$SROCC = 1 - \frac{6}{N(N^2 - 1)} \sum_{i=1}^N d_i^2$$

où d_i est la différence entre les rangs du score objectif et du score subjectif de l'image testée.

- **La fonction objective RMSE**

$$(4.9) \quad f_3 = RMSE(FS, D)$$

$$(4.10) \quad RMSE = \sqrt{\sum_{i=1}^N (FS_i - D_i)^2}$$

FS et D sont les scores objectifs après fusion et les scores subjectifs.

N.B Avant de calculer la fonction objective PLCC, les scores objectifs sont passés (mappés) par une fonction logistique non linéaire comme décrit dans [87].

4.3.2 Étape de fusion

Dans cette deuxième étape, les métriques sélectionnées par l'algorithme PSO sont fusionnées en utilisant le SVR. En effet, ce stage consiste à passer (mapper) des différents scores des métriques à un score final (unique) qui décrit la qualité de l'image. Une description des différentes métriques utilisées dans le processus d'évaluation est donnée par la suite.

4.3.2.1 Métriques de qualité

En plus des cinq métriques décrites dans le chapitre précédent, trois nouvelles métriques ont été ajoutées. En total nous avons huit métriques.

- **Métrique 1** : Blind/Reference-less Image Spatial Quality Evaluator (BRISQUE) [56] (§3.3.1).
- **Métrique 2** : Natural Image Quality Evaluator (NIQE) [57] (§3.3.1).
- **Métrique 3** : Blind Image Quality Index (BIQI) [59] (§3.3.1).
- **Métrique 4** : No-reference Quality Assessment Algorithm for JPEG Compressed Images (NJQA) [26] (§3.3.1).
- **Métrique 5** : No-Reference Quality Assessment of JPEG2000 Compressed Images (JP2KNR) [86] (§3.3.1).
- **Métrique 6** : Blind Image Quality Assessment (BLINDS2) [73] : La Figure 4.2 montre les différentes étapes de cette métrique. Pour chaque image une transformation en cosinus discrète (DCT) est appliquée. Dans cette étape, l'image est décomposée en blocs de taille égale, puis la DCT est calculée localement sur chacun des blocs. Une fois l'extraction des caractéristiques terminée, une densité gaussienne généralisée est utilisée pour modéliser ces caractéristiques. La dernière étape consiste à utiliser un modèle bayésien qui prédit le score final de qualité.

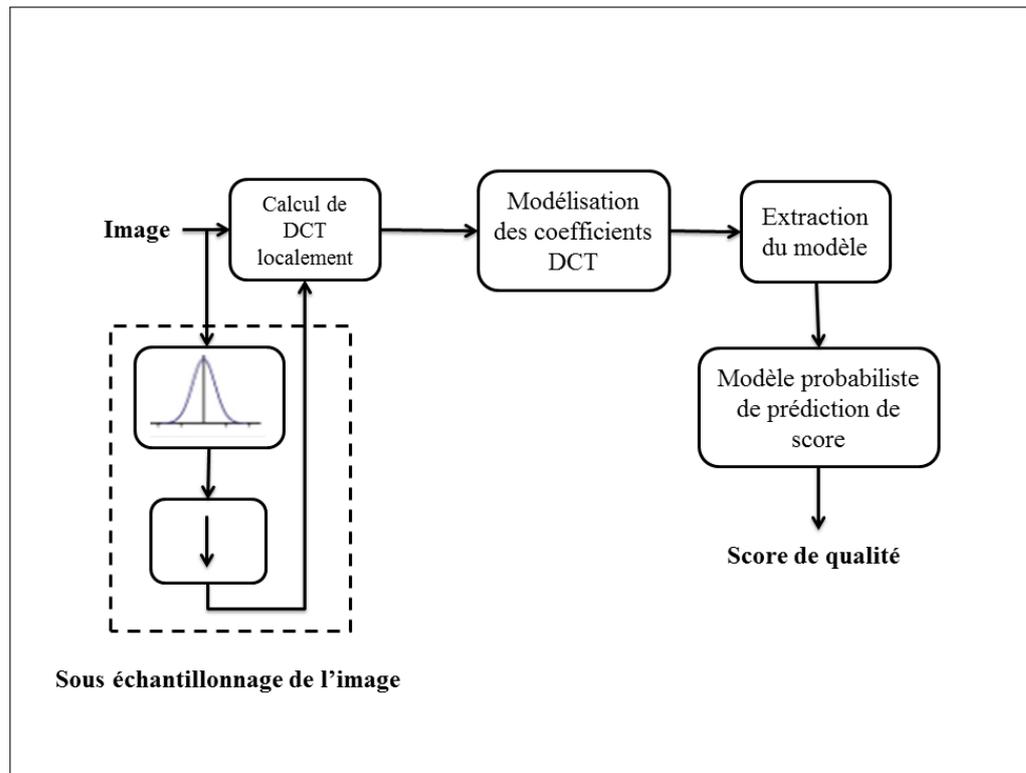


FIGURE 4.2: Synoptique de la métrique 6 (BLINDS2).

- **Métrique 7** : No-Reference Perceptual Quality Assessment of JPEG Compressed Images (JPEGM) [102] : Cette mesure évalue la qualité des images compressées par JPEG. Où le flou et l'effet de bords sont les artefacts les plus significatifs générés au cours du processus de compression. Des caractéristiques ont été extraites, elles peuvent être utilisées pour refléter les grandeurs relatives de ces artefacts. Pour cela, la différence le long de chaque ligne horizontale de l'image est calculée. Ensuite, l'artefact de l'effet du bord est estimé comme étant la différence moyenne entre les limites des blocs. L'activité et la distance moyenne sont combinées pour mesurer le flou. Dans la phase finale et afin d'obtenir le score de qualité, une combinaison de toutes les caractéristiques précédentes a été faite.
- **Métrique 8** : Cumulative Probability of Blur Detection (CPBD) [62] : Cette métrique évalue la qualité des images floues en utilisant la détection de la probabilité cumulative du flou, cumulative probability of blur detection (CPBD). La Figure 4.3 montre ces différentes étapes. Dans un premier temps, l'image est décomposée en blocs. Ensuite, chaque bloc est classé comme un bloc de bord selon un seuil. La largeur des bords est calculée pour chacun des blocs. La prédiction du flou est estimée à l'aide du flou perceptible, just noticeable blur (JNB). Enfin, l'indice de qualité est obtenu en utilisant la probabilité cumulative du flou détecté.

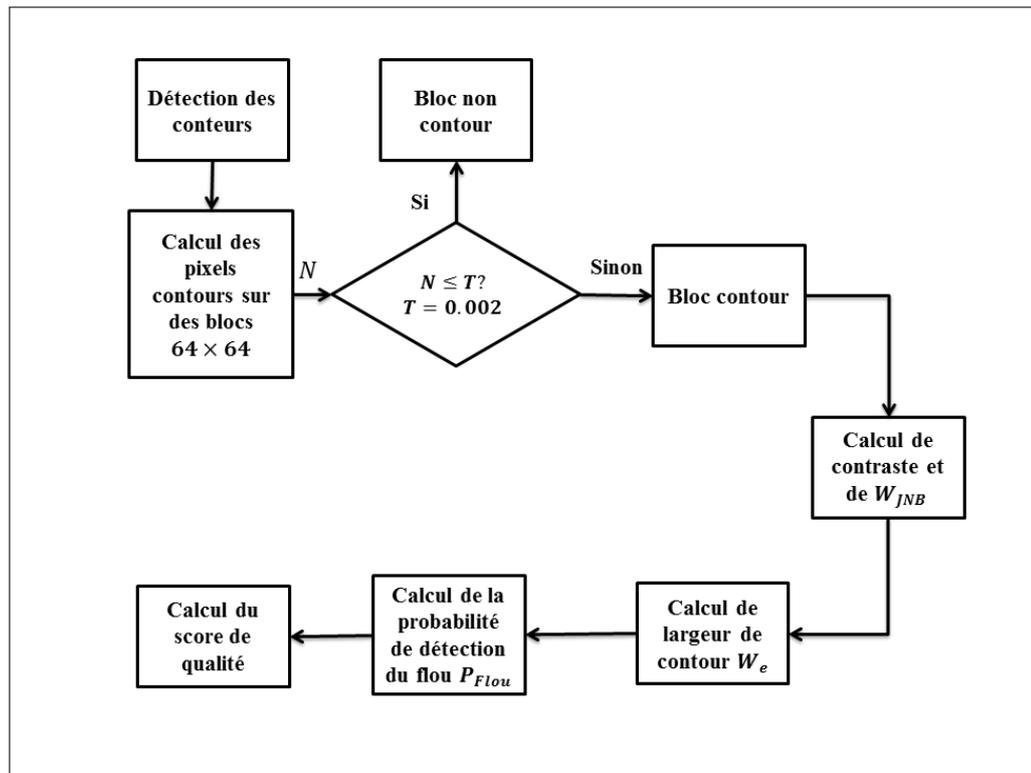


FIGURE 4.3: Synoptique de la métrique 8 (CPBD).

4.3.2.2 Fusion avec les séparateurs à vaste marge de régression (SVR)

Pour fusionner les scores obtenus, la fusion par SVR est adoptée. Une cartographie (mapping) est tirée de l'espace de caractéristiques (différents scores) à un score final de qualité. Le choix du SVR est motivé par sa bonne performance dans les espaces hautement dimensionnés [78].

Le ε -SVR est incorporé dans la métrique proposée de la manière suivante. Les scores des métriques sélectionnées par PSO ainsi que leurs scores subjectifs (DMOS) sont utilisés en entrées du SVR. Pour la i -ème image, utilisant le vecteur de position X_i produit par l'algorithme PSO un vecteur de score est construit comme suit $S_i = (S_1, S_2, \dots, S_M)$, où M est le nombre de métriques. Le score final est prédit en utilisant le modèle obtenu par ε -SVR, qui mappe le vecteur de score à un indice final de qualité.

Pour utiliser le SVR deux étapes sont nécessaires, une pour l'apprentissage et l'autre pour les tests.

- **Étape d'apprentissage** : L'apprentissage est effectué à l'aide d'un ensemble d'images. Pour chaque image, le vecteur de score et le DMOS sont utilisés comme entrées pour le SVR. Un modèle est obtenu, il contient le vecteur de poids w et le biais b . Il est basé sur la minimisation de la différence entre le vecteur de score et les DMOS. Le ε -SVR recherche une différence absolue limitée par ε .

Dans ce travail, le noyau de la fonction de base radiale, radial basis function (RBF), sera

utilisé.

- **Étape de test** : Une fois que le modèle est obtenu, l'étape de test devient facile et on peut évaluer la qualité de toute image. Tout d'abord, le vecteur de score est calculé comme dans l'étape d'apprentissage. Ensuite, il est utilisé avec le modèle obtenu précédemment pour obtenir le score final.

Pour bien clarifier le déroulement de la métrique proposée. Un pseudo-code [1], montrant les différentes étapes de la métrique est présenté ci dessous.

Algorithm 1 PSO-SVR based fusion

SM : Matrice de scores

DMOS : Scores subjectifs

f : Fonction fitness

repeat

Initialisation *X, V*

for *itr* ← 1 to 1000 **do**

S ← *SM* à la position *X*

 ▷ Pour les *X* = 1

 Phase d'apprentissage

S_{train} ← 80% des *S*

 ▷ Pas de chevauchement entre les données d'apprentissage et de test.

DMOS_{train} ← 80% of *DMOS*

Model ← *SVM_{train}* (*S_{train}*, *DMOS_{train}*)

 Phase de test

S_{test} ← 20% of *S*

 ▷ Pas de chevauchement entre les données d'apprentissage et de test.

DMOS_{test} ← 20% of *DMOS*

FS ← *SVM_{predict}* (*S_{test}*, *Model*)

f(itr) ← *fitness*(*FS*, *DMOS_{test}*)

 ▷ ζ peut être une des trois fonctions objectives (f_1, f_2, f_3)

end for

f_{actual} ← *mean*(*f(itr)*)

if *f_{actual}* > *f_{previous}* **then**

f_{previous} ← *f_{actual}*

 Update *V_i*

 Update *X_i*

end if

until Max iterations

On peut dire que les principales contributions de ce travail sont l'adaptation de l'algorithme bio-inspiré PSO dans le domaine de qualité d'image. Ainsi que la fusion de différentes métriques de qualité pour avoir une nouvelle plus robuste.

4.4 Résultats et Discussions

Pour évaluer la performance de la métrique proposée, trois bases d'image ont été utilisées :

- **La base LIVE [87]**
- **La base Tampere (TID2013) [66]**
- **La base CSIQ [40]**

Pour les deux bases d'images TID2013 et CSIQ, les tests ont été limités à quatre dégradations, ceci en raison des métriques utilisées dans le processus de fusion. Ces dégradations sont les suivantes : dégradations dues aux compressions JPEG2000 et JPEG, bruit blanc (WN) et flou Gaussien (GB). En utilisant l'algorithme PSO, différents paramètres ont besoin d'être ajustés. Premièrement, le nombre de particules qui est fixé à 8 (métriques de qualité). Le coefficient d'inertie w est fixé à 0.8. La vitesse prend ces valeurs dans l'intervalle $[-V_{max} V_{max}]$, pour cela les valeurs de la vitesse sont données dans $[-2 2]$. Les deux paramètres α et β doivent répondre à la contrainte suivante $\alpha + \beta \leq 4$, alors ils sont choisis comme suit $\alpha = \beta = 1$. Le nombre d'itération est fixé à 30 itérations pour toutes les expérimentations.

Concernant le SVR, la bibliothèque SVM LIBSVM [9, 33] est utilisée pour l'implémentation de la métrique proposée. Comme nous l'avons vu dans le paragraphe précédent, le SVR a besoins d'un apprentissage qui est nécessaire pour la phase de test. Afin d'assurer un bon apprentissage, chaque base est décomposée aléatoirement en deux sous groupes. Où 80% des images est utilisé pour l'apprentissage et les 20% qui restent pour les tests. Cela est fait pour garantir une meilleure représentation des différentes dégradations. La procédure apprentissage-test est répétée 1000 fois et la moyenne de ces 1000 itérations est donnée.

Le choix de la fonction objective est dicté par le besoin recherché. Si l'on souhaite, par exemple, avoir une corrélation linéaire maximal PLCC, la fonction objective f_1 est utilisée. Il en est de même pour les autres fonctions. En effet, le choix est fixé par le besoin.

Les résultats sont donnés dans les Tableaux 4.1, 4.2 et 4.3. Une comparaison avec l'état de l'art des métriques MSR-EQI est rapportée.

Le Tableau 4.1 contient les résultats de la base LIVE. Les résultats obtenus en terme de corrélation ont démontré clairement que la métrique proposée dépasse les autres métriques. À titre d'exemple, les métriques BRISQUE, NIQE, BIQI et BLINDS2 se comportent mal vis-à-vis la distorsion FF. Par contre, la métrique proposée donne de meilleurs résultats avec des corrélations qui dépassent 0.94.

Tableau 4.1: Résultats obtenus pour la base LIVE

	Métriques \ Dist	JP2K	JPEG	WN	Blur	FF	Toutes
	PLCC	BRISQUE	0.9229	0.9734	0.9851	0.9506	0.9030
NIQE		0,9370	0,9564	0.9773	0.9525	0.9128	0.9147
BIQI		0.8086	0.9011	0.9538	0.8293	0.7328	0.8205
BLINDS2		0.9386	0.9426	0.9635	0.8994	0.8790	0.9146
NJQA		-	0.970	-	-	-	-
JPEGM		-	0.8425	-	-	-	-
JP2KNR		0.92	-	-	-	-	-
CPBD		-	-	-	0.9371	-	-
CORNIA		0.951	0.965	0.987	0.968	0.917	0.935
SFLNIA		0.957	0.958	0.978	0.955	0.920	0.916
Métrique Proposée		0.9608	0.9881	0.9951	0.9806	0.9588	0.9687
SROCC	BRISQUE	0.9139	0.9647	0.9786	0.9511	0.8768	0.9395
	NIQE	0,9172	0,9382	0.9662	0.9341	0.8594	0.9135
	BIQI	0.7995	0.8914	0.9510	0.8463	0.7067	0.8195
	BLINDS2	0.9323	0.9331	0.9463	0.8912	0.8519	0.9124
	NJQA	-	0.967	-	-	-	-
	JPEGM	-	0.6417	-	-	-	-
	JP2KNR	0.94	-	-	-	-	-
	CPBD	-	-	-	0.9449	-	-
	CORNIA	0.943	0.955	0.976	0.969	0.906	0.942
	SFLNIA	0.951	0.947	0.972	0.952	0.911	0.923
	Métrique Proposée	0.9529	0.9783	0.9913	0.9769	0.9462	0.9671
RMSE	BRISQUE	8.484	8.167	4.331	6.635	11.983	11.9845
	NIQE	8.123	6.341	5.321	8.916	15.923	13.4530
	BIQI	14.842	13.7552	8.4094	10.2347	19.2911	15.6223
	BLINDS2	16.614	19.755	13.102	12.058	18.513	17.1282
	NJQA	-	16.3004	-	-	-	-
	JPEGM	-	15.8362	-	-	-	-
	JP2KNR	8.055	-	-	-	-	-
	CPBD	-	-	-	9.5710	-	-
	CORNIA	-	-	-	-	-	-
	SFLNIA	-	-	-	-	-	-
	Métrique Proposée	6.4597	4.5000	2.7513	3.5043	9.4497	6.6543

CHAPITRE 4. ÉVALUATION GÉNÉRIQUE DE LA QUALITÉ BASÉE SUR LES ESSAIS DE PARTICULES ET LES MACHINE À VECTEUR DE SUPPORT

Tableau 4.2: Résultats obtenus pour la base TID2013

	Métrique \ Dist	JP2K	JPEG	WN	Blur	Toutes
PLCC	BRISQUE	0.9169	0.8823	0.8504	0.7965	0.8692
	NIQE	0.9067	0.8904	0.8263	0.8147	0.8075
	BIQI	0.7718	0.8807	0.7962	0.7818	0.7772
	BLINDS2	0.8949	0.8490	0.6451	0.8127	0.7733
	NJQA	-	0.9075	-	-	-
	JPEGM	-	0.8923	-	-	-
	JP2KNR	0.8675	-	-	-	-
	CPBD	-	-	-	0.8576	-
	Métrique Proposée	0.9466	0.9219	0.8831	0.8745	0.9397
SROCC	BRISQUE	0.8925	0.8452	0.8522	0.8142	0.8403
	NIQE	0.8890	0.8434	0.8194	0.7958	0.7962
	BIQI	0.7326	0.8566	0.7842	0.7642	0.7763
	BLINDS2	0.8698	0.8274	0.6400	0.8278	0.7548
	NJQA	-	0.8859	-	-	-
	JPEGM	-	0.8866	-	-	-
	JP2KNR	0.8007	-	-	-	-
	CPBD	-	-	-	0.8518	-
	Métrique Proposée	0.9281	0.9041	0.8682	0.8806	0.8591
RMSE	BRISQUE	0.6795	0.7086	0.3730	0.7544	0.7246
	NIQE	0.7182	0.6854	0.3993	0.7235	0.8276
	BIQI	1.0829	0.7131	0.4289	0.7779	0.8778
	BLINDS2	0.7598	0.7956	0.5417	0.7269	0.7548
	NJQA	-	0.4812	-	-	-
	JPEGM	-	0.4592	-	-	-
	JP2KNR	1.6979	-	-	-	-
	CPBD	-	-	-	0.6420	-
	Métrique Proposée	0.6187	0.3852	0.3149	0.4774	0.6628

Tableau 4.3: Résultats obtenus pour la base CSIQ

	Dist Métrique	JP2K	JPEG	WN	Blur	Toutes
PLCC	BRISQUE	0.8907	0.9346	0.9221	0.9249	0.9190
	NIQE	0.9157	0.8509	0.7661	0.9215	0.8742
	BIQI	0.6944	0.8412	0.8910	0.8470	0.7968
	BLINDS2	0.8841	0.9155	0.7401	0.8319	0.8498
	NJQA	-	0.9332	-	-	-
	JPEGM	-	0.9023	-	-	-
	JP2KNR	0.6343	-	-	-	-
	CPBD	-	-	-	0.8281	-
	Métrique Proposée	0.9302	0.9666	0.9448	0.9381	0.9397
SROCC	BRISQUE	0.8664	0.9040	0.9253	0.9023	0.8989
	NIQE	0.9063	0.8829	0.8097	0.8942	0.8717
	BIQI	0.7084	0.8673	0.8794	0.7712	0.7972
	BLINDS2	0.8847	0.8963	0.7533	0.8492	0.8389
	NJQA	-	0.9247	-	-	-
	JPEGM	-	0.8967	-	-	-
	JP2KNR	0.4319	-	-	-	-
	CPBD	-	-	-	0.8845	-
	Métrique Proposée	0.9065	0.9320	0.9279	0.8936	0.9201
RMSE	BRISQUE	0.1437	0.1088	0.0649	0.1090	0.1114
	NIQE	0.1270	0.1607	0.0980	0.1113	0.1372
	BIQI	0.2274	0.1655	0.0762	0.1524	0.1707
	BLINDS2	0.1476	0.1231	0.1128	0.1590	-
	NJQA	-	0.0920	-	-	-
	JPEGM	-	0.1234	-	-	-
	JP2KNR	0.3131	-	-	-	-
	CPBD	-	-	-	0.1231	-
	Métrique Proposée	0.1052	0.0681	0.0500	0.0931	0.0898

Le Tableau 4.2 donne les résultats obtenus pour la base TID2013. On constate ici que notre métrique donne des résultats prometteurs en les comparant avec les autres métriques. Enfin elle donne les meilleurs résultats pour la dégradation JPEG2000. On peut voir que les valeurs de corrélation sont au dessus de 0.9. Ce qui mets notre métrique en tête par rapport aux autres.

Les résultats obtenus pour la base CSIQ sont rapportés dans le Tableau 4.3. On remarque que les résultats de corrélations dépassent la valeur de 0.9 pour toutes les dégradations, exceptant la dégradation floue.

Ces résultats confirment encore une fois la performance de la méthodes proposée. On note que la métrique proposée a réussi de prédire la qualité perceptuelle pour toutes les distorsions considérées, même si elle intègre certaines métriques dédiées à des distorsions spécifiques. En réalité, elle bénéficie de la spécificité de chaque métrique fusionnée pour augmenter le nombre de dégradations évaluées.

Comme on a précisé avant, le nombre d'itérations pour l'algorithme PSO a été fixé à 30. La Figure 4.4 représente le comportement de la corrélation linéaire PLCC en fonction du nombre d'itérations pour la distorsion due à la compression JPEG2000. On peut remarquer qu'à partir du 20-ième itération la corrélation PLCC se stabilise. Différents tests ont été effectués et enfin le nombre d'itérations à été fixé à 30. On constate également que faire tourner davantage l'algorithme n'a pas amélioré les résultats.

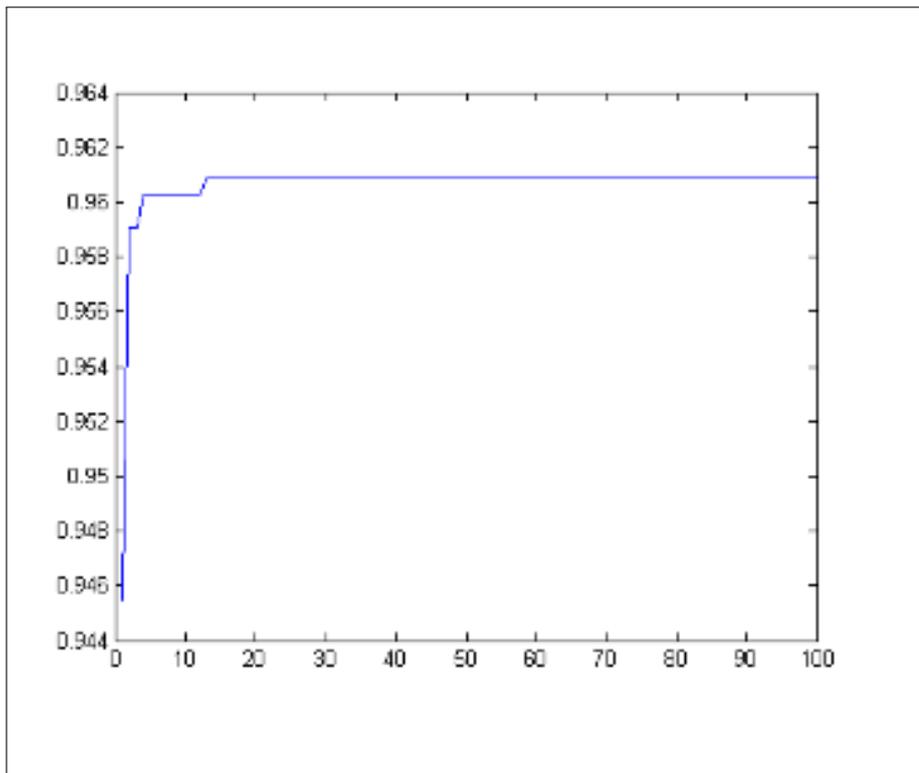


FIGURE 4.4: Courbe du PLCC par rapport au nombre d'itérations de l'algorithme PSO.

La Figure 4.5 présente le nuage des points pour les scores objectifs en fonction des scores subjectifs. L'allure de la courbe indique l'efficacité de la métrique proposée en terme de quantification de la qualité perceptuelle. On remarque également que malgré la présence de différents types de dégradations notre métrique se comporte bien. Tout cela est dû à la stratégie de sélection-fusion adoptée.

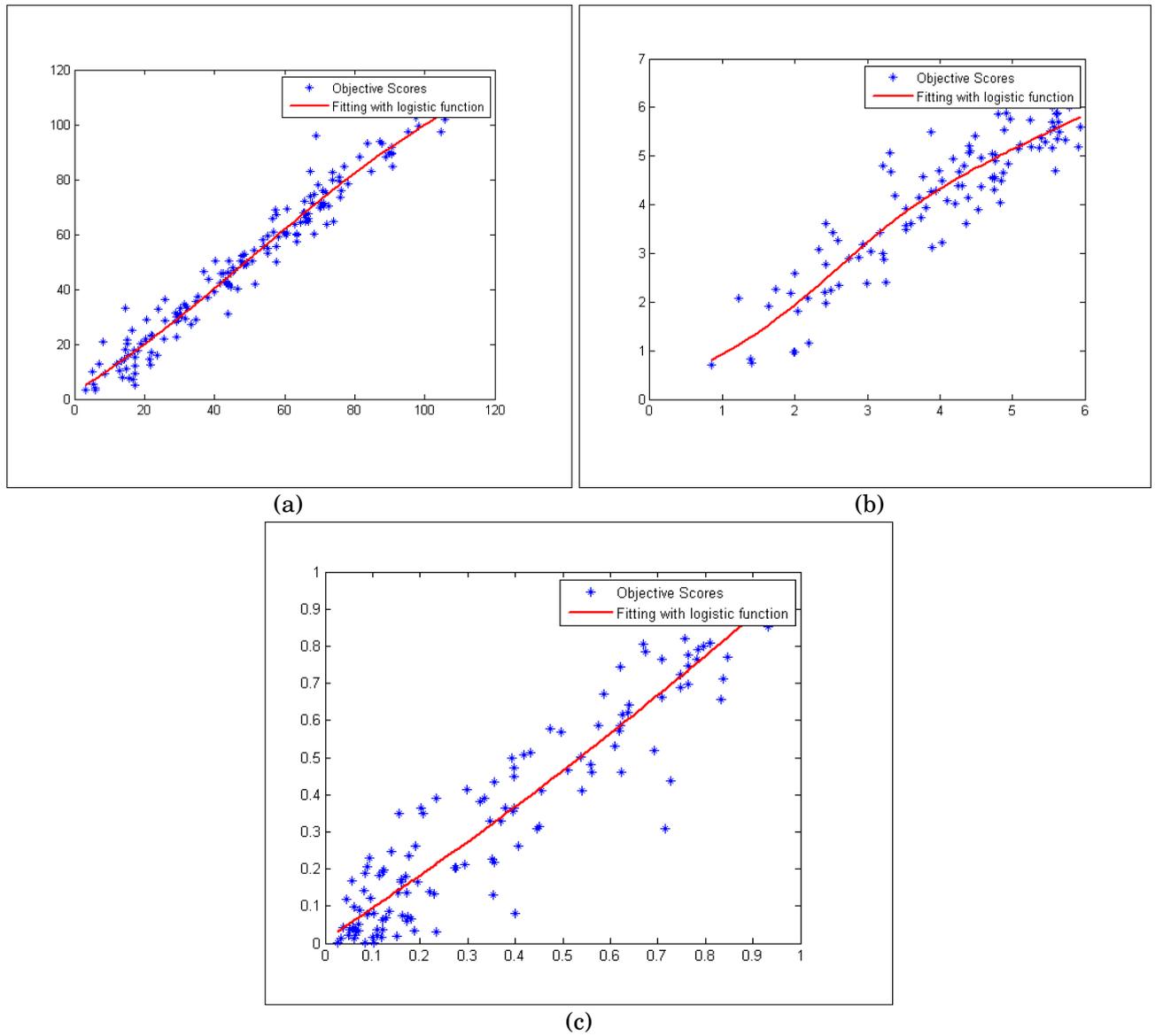


FIGURE 4.5: Tracé des scores objectifs de la métrique proposée en fonction des scores subjectifs : (a) Base LIVE, (b) Base TID2013, (c) Base CSIQ.

Dans le Tableau 4.4, on présente les métriques sélectionnées pour chaque dégradation en utilisant la fonction objective f_2 qui cherche la maximisation de la corrélation de Spearman SROCC. Les tests ont été effectués sur la base LIVE. Vrai indique une bonne sélection. Le taux de bonne sélection est égal à 74.66% pour toutes les dégradations, ceci étant un taux acceptable. Pour la distorsion de transmission FF, on a considéré que les deux métriques JP2KNR et CPBD sont une bonne sélection. Ceci est rendu possible grâce à la distorsion FF produise les mêmes dégradations que le flou et la compression JPEG2000.

Tableau 4.4: Métriques sélectionnées pour chaque dégradation

Métrique \ Dist	JP2K	JPEG	WN	Blur	FF
BRISQUE	Vrai	Vrai	-	Vrai	Vrai
NIQE	-	Vrai	Vrai	Vrai	-
BIQI	-	-	Vrai	Vrai	-
BLINDS2	-	Vrai	-	-	Vrai
NJQA	-	Vrai	Faux	Faux	-
JPEGM	-	Vrai	Faux	Faux	Faux
JP2KNR	Vrai	Faux	-	Faux	Vrai
CPBD	Vrai	-	-	-	Vrai

4.4.1 Indépendance de la base d'apprentissage

Nous avons pu observer dans la partie précédente les différents tests réalisés. Ils ont tous montré la performance de la métrique proposée. Pour valider l'efficacité de notre métrique, on va présenter dans ce paragraphe de nouvelles expérimentations. Elles ont comme objectif la démonstration de l'indépendance des résultats obtenus de la base d'apprentissage. Pour cela, la métrique proposée est entraînée en utilisant la base LIVE et les tests sont effectués sur les deux bases TID2013 et CSIQ. Les résultats obtenus sont présentés dans le Tableau 4.5.

Tableau 4.5: Les corrélations SROCC d'indépendance (Apprentissage sur LIVE et test sur TID2013 et CSIQ)

Base \ Dist	JP2K	JPEG	WN	Blur	Toutes
TID2013	0.8921	0.8449	0.8190	0.8128	0.8456
CSIQ	0.8720	0.9045	0.9060	0.9045	0.9040

En observant les coefficients de corrélation obtenus, on peut dire que la métrique est indépendante de la base d'apprentissage. Si l'on considère les corrélations de la base TID2013, elles sont toutes au dessus de 0.8. On constate la même observation pour la base CSIQ où les valeurs de corrélations sont proches de 0,9 (qui constitue un résultat satisfaisant).

Ceci met en avant l'indépendance de la métrique vis-à-vis des données d'apprentissages. De plus, le fait d'avoir une combinaison de métriques lui donne la performance à évaluer différentes dégradations peu importe la base d'apprentissage.

4.4.2 Comparaison avec les algorithmes génétiques

Dans [65], l'auteur a utilisé les algorithmes génétiques pour la fusion des métriques avec référence MAR-EQI. On a utilisé le même processus pour la fusion des métriques sans référence MSR-EQI afin de comparer cette stratégie avec notre approche. Il a utilisé la fonction objective suivante :

$$(4.11) \quad f = \frac{SROCC(FS,D)}{RMSE(FS,D)}$$

où FS est le score après fusion et D est le score subjectif.

Les bases sont décomposées de la même manière précédente c-à-d deux groupes non chevauchant, où 80% des images ont été utilisées pour l'apprentissage et les 20% restant pour le test.

Tableau 4.6: Comparaison la métrique proposée avec une métrique basée sur les algorithmes génétiques

		JP2K	JPEG	WN	Blur	FF
AG-LIVE	PLCC	0.9487	0.9490	0.8072	0.9645	0.5355
	SROCC	0.9454	0.9568	0.7508	0.9572	0.7308
	RMSE	8.1391	9.4340	18.7705	4.1021	23.4345
Proposée-LIVE	PLCC	0.9608	0.9881	0.9951	0.9806	0.9588
	SROCC	0.9529	0.9783	0.9913	0.9769	0.9462
	RMSE	6.4597	4.5000	2.7513	3.5043	9.4497
AG-TID2013	PLCC	0.9085	0.9452	0.8094	0.9209	-
	SROCC	0.8816	0.9185	0.7504	0.8712	-
	RMSE	0.7198	0.4594	0.3693	0.4926	-
Proposée-TID2013	PLCC	0.9466	0.9219	0.8831	0.8745	-
	SROCC	0.9281	0.9041	0.8682	0.8806	-
	RMSE	0.6187	0.3852	0.3149	0.4774	-
AG-CSIQ	PLCC	0.8293	0.5364	0.7486	0.9411	-
	SROCC	0.9391	0.4915	0.9004	0.8964	-
	RMSE	0.3103	0.7660	0.1094	0.0952	-
Proposée-CSIQ	PLCC	0.9302	0.9666	0.9448	0.9381	-
	SROCC	0.9065	0.9320	0.9279	0.8936	-
	RMSE	0.1052	0.0681	0.0500	0.0931	-

Les résultats obtenus sont rapportés dans le Tableau 4.6. On constate que les résultats obtenus sont acceptables. L'approche basée sur l'algorithme génétiques se comporte bien pour certaines distorsions dans les bases d'images TID2013 et CSIQ (en gras). On peut voir cela pour les distorsions du flou et du JPEG2000, où cette approche dépasse notre métrique. Cependant, en comparant les deux résultats du Tableau 4.6 on observe que la fusion en utilisant notre métrique donne de meilleurs résultats par rapport à ceux de l'algorithme génétique.

4.5 Conclusion

Dans ce chapitre, nous avons présenté une métrique d'évaluation de qualité sans référence MSR-EQI. Elle est basée sur la fusion de différentes métriques. Dans une première étape, le PSO sélectionne les métriques appropriées pour chaque distorsion, puis le SVR est utilisé pour fusionner les scores de ces métriques afin d'obtenir un score final. Il a été démontré que la combinaison des deux algorithmes PSO et SVR pour la conception d'une mesure de qualité est une bonne alternative pour rebondir les difficultés existantes des métriques proposées précédemment. Les tests ont été effectués sur trois bases d'image qui contiennent les distorsions les plus connues. Les résultats obtenus ont démontré la performance de l'approche proposée et le fait qu'elle dépasse les méthodes de l'état de l'art.

**LA THÉORIE DU CONTRÔLE DE GAIN ET L'APPRENTISSAGE
AUTOMATIQUE POUR LA QUALITÉ 3D**

Sommaire

5.1	Introduction	82
5.2	Métriques de l'évaluation des images 3D : État de l'art	82
5.3	Évaluation de la qualité des images 3D	85
5.3.1	Différents modèles de combinaison binoculaire	85
5.3.2	Métrique 3D sans référence basée sur les machines à apprentissage et la théorie du gain	86
5.4	Résultats et discussion	89
5.4.1	Bases d'image 3D	89
5.4.2	Performance de la métrique	89
5.4.3	Expérimentations pour les images Symétriques/Asymétriques	93
5.4.4	Mesure de performance : validation croisée	95
5.4.5	Apprentissage avec différents échantillons	95
5.5	Conclusion	96

5.1 Introduction

De nos jours, le contenu 3D devient très utile dans différents domaines : vidéos mobiles 3D, production des films 3D, outils de capture 3D et de visualisation. Avant la diffusion de ce contenu, il passe par différentes étapes, capture, codage, compression, transmission et enfin visualisation. Chacune de ces étapes peut introduire différents artefacts susceptibles d'affecter la qualité de la perception. En tenant compte du système visuel humain (SVH) très sophistiqué, il est nécessaire de mesurer la qualité de ce contenu avant sa diffusion. À cet effet, la demande d'évaluation de la qualité des images (EQI-3D) est devenue indispensable. Celle-ci est non seulement importante pour la qualité de la perception, mais aussi pour la conception et l'optimisation des systèmes de traitement.

Il est connu que l'évaluation objective peut être classée en trois sous-classes en fonction de la présence de l'image originale (parfaite), évaluation avec référence (MAR-EQI-3D), évaluation avec référence réduite (MRR-EQI-3D) et évaluation sans référence (MSR-EQI-3D). Cependant, pour la qualité d'image 3D, il existe d'autre classification. Dans [11], ils ont regroupé la qualité 3D en deux groupes. Le premier est constitué des métriques qui utilisent des mesures de qualité 2D EQI-2D sans utilisation de l'information profondeur (disparité). En ce qui concerne le deuxième groupe, les métriques 2D sont utilisées avec l'information profondeur pour mieux évaluer la qualité.

Une autre classification est proposée dans [85], où ils ont adopté trois classes. La première concerne l'évaluation basée sur les métriques 2D. Pour la seconde, il s'agit de la qualité qui utilise les informations de la perception 3D. Enfin, la dernière classe rassemble l'évaluation de la qualité 3D sans référence.

La classification suivie dans notre travail est la classification classique : MAR-EQI-3D, MRR-EQI-3D et MSR-EQI-3D.

5.2 Métriques de l'évaluation des images 3D : État de l'art

Dans [113], les auteurs ont utilisé des métriques de qualité d'image 2D, un modèle de seuil différentiel, just noticeable difference (JND), et la disparité pour construire une métrique de qualité 3D. Bensalma et al [5] ont proposé une métrique MAR-EQI-3D pour l'évaluation de la qualité des images stéréoscopiques. Un modèle reproduisant le signal d'énergie binoculaire, généré par des cellules simples et complexes, a été proposé. La différence d'énergie binoculaire entre les stéréo-paires originaux et distordus est utilisée pour produire la qualité perceptuelle. Dans [34], les auteurs ont proposé un indice de qualité d'image basé sur le modèle d'attention visuelle et l'information profondeur. Trois composantes principales ont été étudiées : prédiction de l'attention stéréo, variation de la profondeur et prédiction de distorsion. L'indice de qualité final a été obtenu en combinant toutes ces composantes.

Le travail proposé dans [81], établit une relation entre les valeurs singulières et les scores subjec-

tifs en utilisant la régression SVR. Le score de qualité est alors prédit par le SVR. Benoit et al [4] ont introduit une métrique objective de qualité pour les images stéréoscopiques qui s'appuie à la fois sur l'utilisation de métriques 2D et sur les informations de profondeur pour produire un indice de qualité final. On peut dire que ces métriques sont en général, les extensions des métriques 2D pour la 3D. La plupart d'entre eux ne peuvent pas évaluer les images stéréoscopiques puisqu'elles n'introduisent pas assez d'informations sur la vision binoculaire.

Récemment des métriques MAR-EQI-3D plus efficaces ont été proposées. Dans [120], une métrique basée sur la perception visuelle a été proposée. Tout d'abord, l'image stéréoscopique est décomposée en régions d'occlusion monoculaire et rivalité binoculaire. Des caractéristiques sont extraites à partir de ces deux régions. Dans l'étape finale, les similitudes entre les caractéristiques des stéréo-paires originaux et distordues sont utilisées pour produire le score global de qualité. Lin et al [47] ont utilisé la représentation de fréquences multiples pour traiter l'évaluation de la qualité du contenu 3D. Chaque vue est décomposée en une différence de banc de filtre gaussien (DOG). L'indice de qualité est produit à l'aide d'un processus qui intègre la représentation de fréquences et des métriques 2D.

Dans [83], les auteurs ont proposé une métrique basée sur le concept de réception binoculaire. Les propriétés binoculaires sont utilisées pour simuler des cellules simples et complexes dans le cortex visuel primaire. Premièrement, les dictionnaires de différentes échelles sont utilisés afin de représenter la structure des images. Ensuite, les différences de phase et d'amplitude des vecteurs de coefficients dispersés sont utilisées pour calculer la similarité. La similarité de luminance globale est obtenue à partir des changements de luminance. L'indice de qualité est obtenu en combinant toutes les similarités calculées précédemment.

Dans [43], une métrique MAR-EQI-3D a été proposée. Les similitudes de disparité, de gradient et de congruence de phase ont été utilisées. Ils ont utilisé l'image cyclopéenne pour extraire la congruence de phase et l'amplitude de gradient. Ensuite, les disparités originales et distordues sont calculées. Les similarités résultantes sont utilisées pour produire la qualité finale de l'image stéréoscopique.

Pour les métriques MRR-EQI-3D, quelques approches ont été proposées dans la littérature. Dans [52], les auteurs ont utilisé le tenseur de disparité de couleur pour estimer la carte de disparité des images. Les caractéristiques du système visuel humain sont ensuite intégrées pour créer une fonction de sensibilité de contraste et un seuillage rationnel. Une comparaison entre les coefficients de sensibilité et la cohérence des cartes de disparité permet de produire l'indice de qualité.

Xu et al [107] ont évalué la qualité en mesurant la dégradation structurelle et la compensation parallaxe basée sur la saillance. Premièrement, certaines fonctionnalités sont extraites et envoyées via un canal auxiliaire. Du côté récepteur, une compensation de parallaxe basée sur la saillance et une combinaison non linéaire sont utilisées pour calculer le score de qualité global. Dans [68], une MRR-EQI-3D basée sur des informations perceptives binoculaires a été proposée.

En utilisant les statistiques de distribution des primitives visuelles, l'entropie et l'information mutuelle sont calculées par un modèle de représentation restreint. Ces caractéristiques sont transmises au côté récepteur. La différence entre les caractéristiques originales et déformées est utilisée pour prédire le score de qualité.

Même si quelques mesures MRR-EQI-3D ont été proposées, cela reste toujours insuffisant et en particulier dans le cas où une application en temps réel est nécessaire. En effet, on doit avoir des métriques qui n'ont pas accès à l'information référence. Ce qui nous conduit à la nécessité d'avoir des métriques sans références.

Dans [77], une MSR-EQI-3D a été proposée pour évaluer les images 3D codées par la compression JPEG. Elle est basée sur la disparité et les caractéristiques locales. Des caractéristiques telles que les contours, la texture, l'effet de bord dans chaque bloc de l'image sont utilisés avec la disparité pour prédire la qualité perceptuelle.

Ryu et al [71] ont évalué les distorsions de flou et l'effet de bord dans les images stéréoscopiques. Un modèle pour la perception de la qualité binoculaire est introduit où la relation entre la qualité perceptive et l'information visuelle est étudiée. Tout d'abord, les cartes de flou, de l'effet de bord et de la saillance sont construites pour les images gauche et droite. Puis, les scores de flou et de l'effet de bord sont calculés pour chaque vue respectivement. Enfin, les scores sont combinés en un indice de qualité global à l'aide des modèles de perception binoculaire.

Shao et al [80] ont mis une approche sans référence qui analyse les caractéristiques détaillées des types de distorsions. Ces caractéristiques sont extraites pour décrire la distorsion affectant l'image 3D. La relation entre les caractéristiques stéréoscopiques et les scores subjectifs est établie en utilisant les machine learning SVR, qui ensuite produit l'indice de qualité final.

Chen et al [11] ont utilisé une série de fonctionnalités pour évaluer la qualité des images stéréoscopiques. Une carte de disparité a été générée. Ensuite, un ensemble de réponses de filtre de Gabor multi-échelle est généré. L'image cyclopéenne est alors construite à l'aide des images 3D, de la carte de disparité et des réponses de filtre de Gabor. Des caractéristiques sont calculées à partir de l'image cyclopéenne et de la carte de disparité. Un module d'estimation est utilisé afin de prédire la qualité perceptuelle.

Dans [82] les auteurs ont proposé une métrique sans référence. La prédiction de la qualité est obtenue en deux étapes, entraînement et test. Dans l'étape d'entraînement l'algorithme de représentation parcimonieuse (K-SVD) est utilisé pour former des dictionnaires multi-échelles. Ensuite, les machines learning ont été utilisées pour faire la correspondance entre les vecteurs de caractéristiques extraits des dictionnaires multi-échelle et les notes subjectives correspondantes. L'indice de qualité des paires gauche-droite est obtenu à l'aide du modèle d'apprentissage et le score global est calculé en combinant les scores de l'image gauche et de droite.

Dans [85], une évaluation sans référence pour les images stéréoscopiques a été proposée. Dans la phase d'apprentissage, les auteurs ont obtenu différentes caractéristiques à partir des réponses de l'énergie binoculaire. Lors de l'étape de test, les mêmes caractéristiques sont extraites pour

chaque bloc. Une comparaison a été faite pour produire la qualité optimale.

La mesure proposée est introduite dans la section suivante. Dans laquelle nous avons utilisé des métriques 2D pour évaluer la qualité des différentes sous bandes d'une image filtrée par une série de filtre gaussien, Difference-of-Gaussian (DOG) filter. Les scores obtenus sont ensuite combinés en utilisant les caractéristiques de la perception binoculaire. À la dernière étape, les machines à apprentissage SVR sont utilisées pour calculer le score final de qualité.

5.3 Évaluation de la qualité des images 3D

Pour évaluer la qualité des images stéréoscopiques une nouvelle approche est proposée. Elle est basée sur la fusion des scores avec l'intégration des caractéristiques de la perception binoculaire. Tout d'abord, les deux vues gauche-droite sont décomposées puis filtrées par des filtres de Différence Gaussienne (DOG). Ensuite, la théorie du gain est utilisée pour pondérer les scores des MSR-EQI-2D appliqués à chaque bande gauche-droite. Dans l'étape finale, l'apprentissage automatique est utilisé pour produire le score final. Dans le paragraphe suivant on va présenter les différentes techniques utilisées pour modéliser la combinaison binoculaire du système visuelle humain.

5.3.1 Différents modèles de combinaison binoculaire

Une meilleure compréhension des études physiologiques de la vision binoculaire est indispensable pour le développement de nouvelles métriques EQI-3D. Pour cela, différents modèles de combinaison binoculaires ont été proposés dans la littérature. Parmi ces modèles, on a la combinaison binoculaire de la luminosité ou du contraste [3, 18, 27]. La perception binoculaire de luminosité pour une image stéréoscopique peut être donnée par la fonction f_B de la luminosité des vues gauche et droite comme suit :

$$(5.1) \quad I_{Binoc} = f_B(I_G, I_D)$$

avec I_{Binoc} l'image perçue après combinaison et I_G, I_D les deux vues gauche-droite respectivement. La fonction de combinaison f_B est exprimée à l'aide de différents modèles, pondération des vues, sommation quadratique et sommation de vecteur.

- **Modèle de pondération des vues** : ce modèle décrit la combinaison binoculaire comme la somme pondérée de l'image gauche et droite. La fonction f_B est donnée par l'équation 5.2

$$(5.2) \quad f_B = w_1 \cdot I_G + w_2 \cdot I_D$$

avec w_1, w_2 sont les poids de l'image gauche et droite respectivement.

- **Modèle de sommation quadratique** : pour ce modèle la combinaison binoculaire est exprimée par la sommation quadratique des deux luminances gauche-droite comme suit :

$$(5.3) \quad f_B = \sqrt{I_G^2 + I_D^2}$$

Ce modèle est valable pour des luminances symétriques pour les deux yeux.

- **Modèle de sommation de vecteur** : ce modèle est basé sur les signaux de flux de luminosité monoculaire de l'œil gauche et de l'œil droit comme suit :

$$(5.4) \quad f_B = \frac{\epsilon_G^2 + \epsilon_D^2}{\epsilon_G + \epsilon_D}$$

avec (ϵ_G, ϵ_D) les flux de signaux des deux images gauche et droite.

- **Modèle des réseaux de neurones** : il utilise les réponses neuronales de chaque œil pour le calcul de la combinaison binoculaire.

$$(5.5) \quad f_B = N_G + N_D$$

avec N_G, N_D les réponses neuronales des yeux gauche et droit respectivement.

- **Modèle de théorie de gain-contrôle** : dans [18] les auteurs ont proposé une étude pour la combinaison binoculaire basée sur deux principes d'interaction. Pour chaque voisinage spatial, chaque œil exerce un contrôle de gain sur le signal de l'autre œil en proportion de l'énergie de contraste. Ce dernier exerce aussi un contrôle de gain sur celui de l'autre œil. En tenant compte de ces spécificités, un simple modèle physiologique qui décrit la combinaison binoculaire a été proposé comme suit :

$$\begin{aligned} f_B &= \left(\frac{1 + Eng_G}{1 + Eng_G + Eng_D} \right) I_G + \left(\frac{1 + Eng_D}{1 + Eng_G + Eng_D} \right) I_D \\ &= G_G \cdot I_G + G_D \cdot I_D \end{aligned}$$

(5.6)

avec Eng_G, Eng_D les sommes d'énergie sur l'ensemble des fréquences pour l'œil gauche et l'œil droit respectivement.

Le modèle de "théorie de gain-contrôle" est utilisé dans notre nouvelle métrique d'évaluation de qualité d'image 3D.

5.3.2 Métrique 3D sans référence basée sur les machines à apprentissage et la théorie du gain

Plusieurs études psychophysiques ont pu démontrer que la réponse visuelle du SVH peut être approximée par une décomposition en sous bandes. Et cela en utilisant une série de différence de filtre gaussien DOG [104, 105]. La décomposition est donnée comme suit :

$$(5.7) \quad I_f = (G(\sigma) - G(k\sigma)) * I$$

avec σ l'écart-type du filtre gaussien, I représente soit l'image de gauche ou de droite, et k une constante qui est égale à 1.6 [104].

Pour générer les différents filtres on a utilisé plusieurs valeurs de σ , elles sont données par :

$$(5.8) \quad \sigma_i = \begin{cases} 0, & \text{si } i = 0 \\ 1, & \text{si } i = 1 \\ k\sigma_{i-1}, & \text{si } i > 1 \end{cases}$$

Les réponses des filtres DOG peuvent être représentées comme suit :

$$(5.9) \quad \begin{aligned} DOG(I) &= [DOG(I;\sigma_0), DOG(I;\sigma_1), \dots, DOG(I;\sigma_N)] \\ &= [DOG_0, DOG_1, \dots, DOG_N] \end{aligned}$$

avec N le nombre de sous bandes qui peut être atteint.

En utilisant les filtres DOG différentes réponses ont été générées pour chaque vue de l'image stéréoscopique.

$$(5.10) \quad DOG(I_G) = [DOG_0^G, DOG_1^G, \dots, DOG_N^G]$$

et

$$(5.11) \quad DOG(I_D) = [DOG_0^D, DOG_1^D, \dots, DOG_N^D]$$

avec $DOG(I_G), DOG(I_D)$ indiquent les réponses de chaque DOG et N le nombre de sous bandes. Le modèle de théorie de gain-contrôle extrait des différentes réponses des filtres DOG est calculé comme suit :

$$(5.12) \quad G_i^G = \frac{1 + Eng(DOG_i^G)}{1 + Eng_G + Eng_D}$$

avec $Eng(DOG_i^G)$, Eng_G et Eng_D :

$$(5.13) \quad Eng(DOG_i^G) = \sum_{p \in DOG_i^G} p^2$$

$$(5.14) \quad Eng_G = \sum_{j=1}^N (Eng(DOG_j^G))$$

$$(5.15) \quad Eng_D = \sum_{j=1}^N (Eng(DOG_j^D))$$

avec p les pixels dans DOG_i^G et N le nombre de bandes.

La quantité $Eng(DOG_i^l)$ est appelée l'énergie fréquentielle de la bande. En ce qui concerne le modèle de théorie de gain-contrôle pour l'image droite les mêmes étapes sont à suivre.

5.3.2.1 Combinaison des scores

Une fois le modèle de combinaison binoculaire est calculé comme on a vu précédemment. Différentes métriques 2D sont utilisées pour évaluer la qualité perceptuelle des différentes bandes. Ensuite, le modèle de théorie de gain-contrôle est utilisé pour pondérer les scores 2D comme le montre l'équation 5.16

$$(5.16) \quad s_i(I_G, I_D) = G_i^G \cdot 2DM(DOG_i^G) + G_i^D \cdot 2DM(DOG_i^D)$$

avec s le résultat de combinaison des deux scores de la bande gauche-droite i , $2DM$ est le score 2D, et (DOG_i^G, DOG_i^D) sont les bandes gauche-droite du filtre DOG.

Les mesures 2D MSR-EQI utilisées sont décrites précédemment et ce sont : BRISQUE (§3.3.1), NIQE (§3.3.1), BIQI (§3.3.1), JP2KNR (§3.3.1) et JPEGM (§4.3.2.1).

Au final, une matrice de scores S est obtenue comme suit :

$$S = \begin{bmatrix} s_{11} & s_{12} & s_{13} & \dots & s_{1M} \\ s_{21} & s_{22} & s_{23} & \dots & s_{2M} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ s_{N1} & s_{N2} & s_{N3} & \dots & s_{NM} \end{bmatrix}$$

M, N représentent le nombre de bandes DOG et le nombre de métriques 2D utilisés respectivement.

5.3.2.2 Prédiction de score de qualité

Le score final de qualité est prédit en utilisant le SVR [78, 88] comme dans le paragraphe 3.3.2. Nous avons besoins pour chaque image stéréoscopique d'apprentissage de la matrice de scores S et le score objectif correspondant. L'équation (3.19) permet d'obtenir le modèle d'apprentissage. Pour évaluer la qualité d'une nouvelle image on fait appel à l'équation (3.20) pour prédire l'indice de qualité et cela en se basant sur le modèle d'apprentissage.

5.4 Résultats et discussion

5.4.1 Bases d'image 3D

Il existe différentes bases d'images de qualité 3D. On a utilisé parmi elles les plus connues :

- **La base LIVE 3D phase I [61].**
- **La base LIVE 3D phase II [11].**
- **La base Waterloo-IVC 3D Phase I [96].**
- **La base Waterloo-IVC 3D phase II [95].**

5.4.2 Performance de la métrique

Pour comparer la métrique proposée avec les métriques de l'état de l'art, les deux corrélations linéaire (PLCC) et de Spearman (SROCC) ont été utilisées. Elles sont calculées entre les scores objectifs et subjectifs. Des PLCC et SROCC proche de 1 indiquent la bonne performance de l'évaluation.

Pour les expérimentations, des répartitions de chaque base d'image ont été faites. Pour cela, 80% des images sont utilisées pour l'apprentissage et les 20% restantes pour les tests.

Les résultats de corrélations PLCC et SROCC pour la base LIVE 3D phase I sont indiqués respectivement dans les Tableaux 5.1 et 5.2.

Tableau 5.1: Les résultats de corrélation PLCC (LIVE 3D Phase I)

Métrique \ Dist	JP2k	JPEG	WN	Blur	FF
FIPSNR [47]	0.8381	0.2866	0.9280	0.9475	0.7086
FISSIM [47]	0.8210	0.2741	0.9250	0.9080	0.7297
Benoit [4]	0.8859	0.5766	0.9354	0.9217	0.7477
Bensalma [5]	0.8389	0.3803	0.9147	0.9369	0.7339
Chen [12]	0.9164	0.6344	0.9436	0.9417	0.7580
Shao [84]	0.9238	0.6563	0.9410	0.9513	0.8403
Ref [83]	0.9213	0.5200	0.9448	0.9592	0.8594
Ref [82]	0.9495	0.7959	0.9378	0.9856	0.8368
Ref [85]	0.8723	0.5972	0.9161	0.9233	-
Proposée	0.9471	0.7583	0.9667	0.9724	0.8855

Tableau 5.2: Les résultats de corrélation SROCC (LIVE 3D Phase I)

Métrique \ Dist	JP2k	JPEG	WN	Blur	FF
FIPSNR [47]	0.8388	0.2070	0.9284	0.9345	0.6581
FISSIM [47]	0.8222	0.2407	0.9282	0.8788	0.6866
Benoit [4]	0.8730	0.4983	0.9369	0.8802	0.6342
Bensalma [5]	0.8171	0.3283	0.9055	0.9157	0.6500
Chen [12]	0.8956	0.5582	0.9481	0.9261	0.6879
Shao [84]	0.8752	0.6148	0.9431	0.9375	0.8714
Ref [83]	0.7845	0.7330	0.9651	0.9204	0.8905
Ref [82]	0.9363	0.8182	0.9348	0.9273	0.8142
Ref [85]	0.9003	0.6073	0.9265	0.9235	-
Proposée	0.8945	0.7412	0.9496	0.9383	0.8230

Notre métrique dépasse la plupart des méthodes de l'état de l'art. En effet, seules quelques mesures donnent de meilleurs résultats pour certaines distorsions. Cependant, ces derniers sont essentiellement des métriques avec référence MAR-EQI-3D. La métrique proposée dans [82], qui est une mesure sans référence, dépasse notre métrique où elle a une légère différence pour les distorsions JP2K et JPEG.

Les Tableaux 5.3 et 5.4 présentent les résultats obtenus pour la base LIVE 3D phase II. On peut remarquer que la métrique proposée est très compétitive vis-à-vis des métriques de l'état de l'art.

Tableau 5.3: Les résultats de corrélation PLCC (LIVE 3D Phase II)

Métrique \ Dist	JP2k	JPEG	WN	Blur	FF
FIPSNR [47]	0.7457	0.6124	0.9150	0.7083	0.7025
FISSIM [47]	0.7191	0.5486	0.9139	0.7250	0.7342
Benoit [4]	0.6467	0.5328	0.8610	0.8814	0.8472
Bensalma [5]	0.6667	0.8577	0.9436	0.9077	0.9097
Chen [12]	0.8426	0.8422	0.9602	0.9650	0.9097
Shao [84]	0.8377	0.7504	0.8496	0.8270	0.8808
Ref [83]	0.7823	0.7472	0.9464	0.9580	0.9046
Proposée	0.9530	0.8738	0.9499	0.9767	0.9399

Tableau 5.4: Les résultats de corrélation SROCC (LIVE 3D Phase II)

Métrique \ Dist	JP2k	JPEG	WN	Blur	FF
FIPSNR [47]	0.7193	0.6129	0.9073	0.7112	0.7012
FISSIM [47]	0.7003	0.5644	0.9091	0.7387	0.7350
Benoit [4]	0.6325	0.5078	0.8569	0.8545	0.8319
Bensalma [5]	0.8038	0.8461	0.9386	0.8838	0.8743
Chen [12]	0.8334	0.8396	0.9554	0.9096	0.8890
Shao [84]	0.8477	0.7195	0.8455	0.8005	0.8509
Ref [83]	0.7845	0.7330	0.9651	0.9204	0.8904
Proposée	0.9212	0.8294	0.9120	0.9084	0.9070

Dans le tableau 5.5, on présente les corrélations PLCC et SROCC obtenus à partir des tests conduits sur la base Waterloo-IVC 3D phase I. Les images stéréoscopiques ici sont regroupées en quatre groupes : WN, Blur, JPEG et Autres. La catégorie "Autres" contient les images avec différentes distorsions.

Il est clair que notre métrique a des corrélations qui coïncident avec les jugements humains. Cela est vérifié même pour le groupe "Autres" qui contient différents types et niveaux de distorsions.

Tableau 5.5: Résultats de la métrique proposée sur la base Waterloo-IVC 3D phase I

Indice \ Dist	WN	Blur	JPEG	Autres
PLCC	0.9379	0.9806	0.9635	0.9381
SROCC	0.8393	0.9154	0.8693	0.9062

Les mêmes tests ont été conduits sur la base Waterloo-IVC 3D phase II et les résultats obtenus sont donnés dans le tableau 5.6. Pour cette base, nous avons eu des coefficients de corrélation qui sont supérieurs à 0.91, ce qui montre l'efficacité de la nouvelle métrique.

Tableau 5.6: Résultats de la métrique proposée sur la base Waterloo-IVC 3D phase II

Indice \ Dist	WN	Blur	JPEG	Autres
PLCC	0.9513	0.9420	0.9302	0.9303
SROCC	0.8961	0.8631	0.8377	0.9208

Pour montrer l'allure des scores objective vis-à-vis des scores subjectifs, on a ploté ces deux derniers et le résultat est visible dans la Figure 5.1. On constate qu'il existe dans cette figure, une forte cohérence entre les scores de notre métrique et les scores subjectifs.

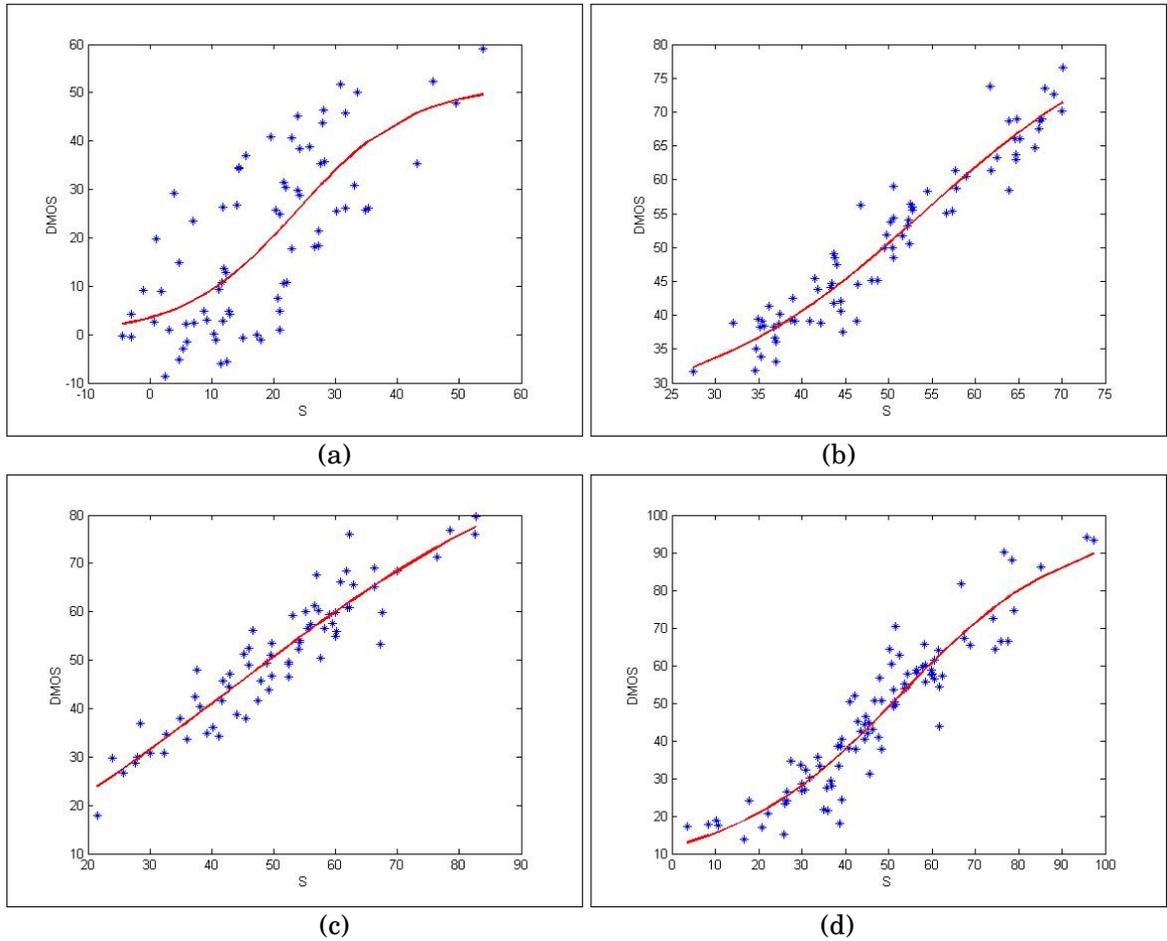


FIGURE 5.1: Plot des scores de la métrique proposée et les scores subjectifs : (a) LIVE 3D phase I, (b) LIVE 3D phase II, (c) Waterloo-IVC 3D phase I, (d) Waterloo-IVC 3D phase II

5.4.3 Expérimentations pour les images Symétriques/Asymétriques

Dans cette partie, il est question d'étudier le comportement de la métrique vis-à-vis des distorsions symétriques et asymétriques. La symétrie dans une image stéréoscopiques signifie que les deux vues ont été dégradées par le même type de distorsion et vice-versa. Les trois bases LIVE 3D phase II, Waterloo-IVC 3D phase I et Waterloo-IVC 3D phase II contiennent les deux types de déformations. Les tests ont été fait en décomposant les base en deux groupes symétrique et asymétrique. Le Tableau 5.7 montre les coefficients de corrélation de notre métrique ainsi qu'une comparaison avec les mesures de l'état de l'art.

Tableau 5.7: Résultats pour les images symétrique et asymétrique

		Waterloo-IVC 3D phase I			Waterloo-IVC 3D phase II			LIVE 3D phase II		
Métrique		Sym	Asym	Toutes	Sym	Asym	Toutes	Sym	Asym	Toutes
PLCC	ID-SSIM [95]	0.9638	0.9286	0.9300	0.9377	0.8799	0.8918	0.9372	0.8981	0.9159
	You [113]	0.8681	0.7089	0.7125	0.7634	0.6857	0.6817	0.8245	0.7832	0.8015
	Benoit [4]	0.8503	0.6970	0.6797	0.7549	0.5548	0.5507	0.7339	0.7701	0.7624
	Yang [110]	0.8359	0.7150	0.7061	0.7920	0.6413	0.6388	0.7752	0.7088	0.7346
	Chen [12]	0.9553	0.7324	0.7337	0.8371	0.6330	0.6130	0.9384	0.8753	0.9073
	Proposée	0.9478	0.9375	0.9378	0.9236	0.9315	0.9311	0.9387	0.9224	0.9314
SROCC	ID-SSIM [95]	0.9481	0.9097	0.9177	0.9052	0.8482	0.8687	0.9234	0.9019	0.9188
	You [113]	0.7517	0.5706	0.5968	0.5602	0.5997	0.5873	0.8030	0.7721	0.7924
	Benoit [4]	0.7275	0.5766	0.5852	0.5713	0.4539	0.4595	0.6959	0.7474	0.7436
	Yang [110]	0.6668	0.6108	0.6106	0.6627	0.5946	0.5875	0.7608	0.6960	0.7210
	Chen [12]	0.9241	0.6428	0.6815	0.7581	0.5627	0.5781	0.9252	0.8538	0.9013
	Proposée	0.9158	0.9063	0.9178	0.9000	0.9219	0.9216	0.9043	0.9057	0.9236

A partir de l'étude du Tableau 5.7, on constate que peu de métriques de l'état de l'art ont effectué ce type de test. Cela est due simplement à la difficulté et à la dureté de ce dernier. En l'occurrence, notre métrique était très compétitive, elle a dépassé toutes les autres mesures. Une seule métrique ID-SSIM a donné de meilleurs résultats. Néanmoins, cela n'influence pas nos résultats car cette dernière reste une métrique avec référence. Une explication simple de ces résultats peut être donnée par l'utilisation du gain de contrôle exercé par chaque vue et les différentes métriques d'évaluation 2D.

La Figure 5.2 présente le nuage de points des scores prédits par rapport aux scores objectifs. Les points sont plus proches de la ligne rouge qui représente les scores objectifs. Cela indique que la métrique coïncide bien avec les jugements humains. On peut dire que la robustesse de la métrique est garantie peu importe la nature de la dégradation.

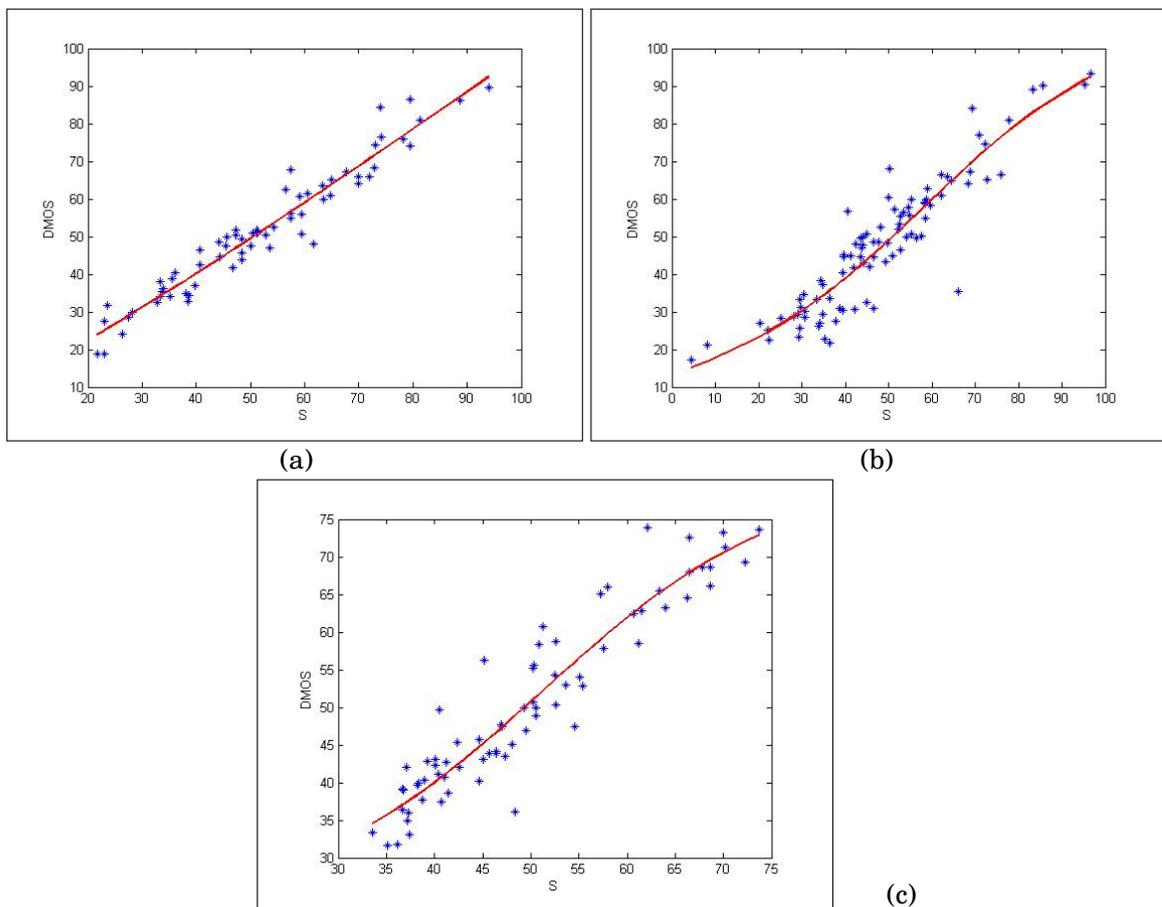


FIGURE 5.2: Plot des scores objectifs pour les images Sym-Asym avec les DMOS correspondant : (a) Waterloo-IVC 3D phase I, (b) Waterloo-IVC 3D phase II, (c) LIVE 3D phase II

5.4.4 Mesure de performance : validation croisée

Afin de généraliser les résultats obtenus et approuver qu'ils sont indépendants de la base d'apprentissage, on a effectué une nouvelle expérimentation qui est la validation croisée. Pour cela, l'apprentissage est fait sur la base LIVE 3D phase II. Une fois le modèle construit, on l'utilise pour évaluer les différentes bases (Waterloo-IVC 3D phase I, Waterloo-IVC 3D phase II, LIVE 3D phase I). Le Tableau 5.8 présente les résultats obtenus pour chaque base.

Tableau 5.8: La performance en utilisant le modèle d'apprentissage de la base LIVE 3D phase II.

	Waterloo-IVC 3D phase I		Waterloo-IVC 3D phase II		LIVE 3D phase I	
	PLCC	SROCC	PLCC	SROCC	PLCC	SROCC
WN	0.9269	0.9330	0.8949	0.9006	0.8843	0.7998
Blur	0.9431	0.9217	0.9324	0.9105	0.8569	0.8275
JPEG	0.8495	0.8600	0.7358	0.7137	0.5659	0.5587
JP2K	-	-	-	-	0.9027	0.8666
FF	-	-	-	-	0.6380	0.7792

On utilisant le modèle obtenu de la base LIVE 3D phase II, les coefficients de corrélation indiquent bien que la mesure est indépendante de la base d'apprentissage. On peut signaler qu'il y avait des difficultés pour les distorsions JPEG et FF en ce qui concerne la base de LIVE 3D phase I. Cela est due à la base elle-même. Pour assurer l'indépendance de la métrique, il faut avoir des données d'apprentissage représentatives des différentes distorsions.

5.4.5 Apprentissage avec différents échantillons

Comme mentionné auparavant, la nouvelle métrique utilise les machines learning basées sur l'apprentissage. Pour évaluer la robustesse de la métrique vis-à-vis des échantillons d'apprentissage, différents échantillons ont été utilisés. Ce sont comme suit : 80%, 50% et 30% des images pour l'apprentissage et le reste pour le test. La même procédure de répétition a été suivie pour l'apprentissage-test (1000 fois). La moyenne des corrélations PLCC et SROCC est donnée dans le Tableau 5.9.

Tableau 5.9: Moyenne de PLCC et SROCC pour différents groupes d'apprentissage-test

		Waterloo-IVC 3D phase I			Waterloo-IVC 3D phase II			LIVE 3D phase II		
	Échantillons	Sym	Asym	All	Sym	Asym	All	Sym	Asym	All
PLCC	80%	0.9478	0.9375	0.9378	0.9236	0.9315	0.9311	0.9387	0.9224	0.9314
	50%	0.9249	0.9134	0.9147	0.8936	0.8826	0.8888	0.9250	0.9073	0.9175
	30%	0.8905	0.8564	0.8645	0.8406	0.8039	0.8186	0.9053	0.8814	0.8948
SROCC	80%	0.9158	0.9063	0.9178	0.9000	0.9219	0.9216	0.9043	0.9057	0.9236
	50%	0.9154	0.8910	0.9021	0.8870	0.8765	0.8826	0.9018	0.8990	0.9160
	30%	0.8844	0.8418	0.8578	0.8354	0.7964	0.8127	0.8820	0.8736	0.8945

Dans le cas de 80%, la métrique a donné de meilleurs résultats pour toutes les bases, où les coefficients de corrélation dépassent la valeurs de 0,9. Lorsqu'on a utilisé 50% des images pour l'apprentissage, les résultats obtenus ne sont pas comme dans le premier cas, mais ils sont prometteurs. On a pu avoir des coefficients de corrélations autour de 0,9. Dans le cas de 30% des échantillons, la performance de la métrique proposée a chuté par rapport aux deux cas précédents.

De nouveau, on peut dire que la métrique proposée est robuste. Dans tout les cas, il faut s'assurer que les images d'apprentissage représentent les différentes distorsions à évaluer.

5.5 Conclusion

Dans ce chapitre, une nouvelle métrique sans référence a été proposée pour évaluer la qualité des images stéréoscopiques. Elle est différente des métriques EQI 3D basées sur les mesures 2D, notre méthode intègre en plus les caractéristiques du système visuel humain, en particulier la combinaison binoculaire. Pour cela, cinq MSR-EQI 2D ont été utilisées pour évaluer la qualité des différentes bandes de décomposition DOG. Ensuite, la théorie de contrôle du gain simulant la vision binoculaire humaine a été utilisée pour combiner les scores. Afin de calculer le score final, on a utilisé les machines learning SVR. La performance de la métrique proposée a été évalué en utilisant différentes bases d'images. Les prédictions de la métrique proposée ont montré une forte corrélation par rapport aux scores subjectifs.

CONCLUSION GÉNÉRALE

Le travail entre nos mains présente une étude dans laquelle le sujet de l'évaluation de qualité d'image est traité. Vu l'importance de l'évaluation pour la recherche ou bien pour les applications temps réel, nous nous sommes intéressés l'évaluation sans référence. La majorité des métriques existantes dans la littérature sont avec spécifications de dégradation, où elle est bien connue au préalable. Mais en réalité une image peut contenir plusieurs distorsions en même temps. Une image compressée par la compression JPEG2000 est un meilleurs exemple, elle peut avoir des oscillations autour des contours (ringing) comme elle peut avoir aussi la dégradation flou. De cela, il était essentiel pour nous d'évaluer la qualité en prenant en considération toutes les distorsions susceptibles d'être présentes. Les travaux précédents ont bien montré qu'une seule métrique n'est pas capable d'assurer une bonne évaluation. L'idée c'était d'avoir une métrique la plus générique possible.

La fusion des métriques représente l'une des prometteuses solutions, utilisé les plus efficace dans la littérature avec un système de fusion pour assurer une meilleures estimation en présence de plusieurs distorsions. Les machines à apprentissage tel que les machines à vecteur de support SVR représente un outil puissant, capable de produire à partir des scores objectifs et subjectifs un score unique qui décrit la qualité de l'image. Afin de fusionner les métriques les plus adaptées, un système de sélection inspiré de la nature a été utilisé.

La même idée est utilisée pour mesurer la qualité des images stéréoscopiques. Cette fois-ci, les caractéristiques du système visuel humain sont introduites. La théorie du gain pour la vision binoculaire est utilisée pour simuler la combinaison effectuée par le SVH.

Cette étude a bien montré qu'il est possible de proposer un système générique d'évaluation. Les résultats obtenus à travers les différentes expérimentations sont témoins de l'efficacité et la performance de notre modèle. La spécificité de notre système c'est qu'il peut être toujours amélioré et optimisé. Cela est garantie par l'ajout de nouvelles métriques.

Comme perspective nous proposons, de poursuivre la recherche dans les domaines suivants :

Qualité des images à grande gamme dynamique (high dynamic range HDR) Les algorithmes de tone-mapping (TMO) ou bien les algorithmes de conversion vers les images à faible dynamique (de HDR à low dynamic range LDR) fournissent un outil utiles pour la visualisation des images HDR sur les écrans LDR standard. Différents TMO créent différentes images, et

une question naturelle c'est laquelle qui a la meilleure qualité. Sans une mesure de qualité appropriée, il est difficile de comparer les différents algorithmes TMO. C'est un axe de recherche qui est ouvert.

Les images médicales L'objectif de l'imagerie médicale est d'aider les praticiens de la santé à rendre des diagnostics précis. Ce qui nécessite des images avec une qualité optimale afin d'assurer les bonnes décisions. Devant cette incontournable importance, l'imagerie médicale est devenue présente dans la prise en charge des patients que ce soit au cours du dépistage ou dans la phase de diagnostique. En effet, l'évaluation de la qualité perceptuelle est indispensable pour la comparaison, voire l'optimisation des systèmes d'imagerie. Ce qui représente un axe de recherche à développer.

BIBLIOGRAPHIE

- [1] P. K. ATREY, M. A. HOSSAIN, A. EL SADDIK, AND M. S. KANKANHALLI, *Multimodal fusion for multimedia analysis : a survey*, *Multimedia systems*, 16 (2010), pp. 345–379.
- [2] R. V. BABU AND A. PERKIS, *An hvs-based no-reference perceptual quality assessment of jpeg coded images using neural networks*, in *IEEE International Conference on Image Processing 2005*, vol. 1, Sept 2005, pp. I–433–6.
- [3] D. H. BAKER, S. A. WALLIS, M. A. GEORGESON, AND T. S. MEESE, *Nonlinearities in the binocular combination of luminance and contrast*, *Vision research*, 56 (2012), pp. 1–9.
- [4] A. BENOIT, P. LE CALLET, P. CAMPISI, AND R. COUSSEAU, *Quality assessment of stereoscopic images*, *EURASIP journal on image and video processing*, 2008 (2009), pp. 1–13.
- [5] R. BENSALMA AND M.-C. LARABI, *A perceptual metric for stereoscopic image quality assessment based on the binocular energy*, *Multidimensional Systems and Signal Processing*, 24 (2013), pp. 281–316.
- [6] B. E. BOSER, I. M. GUYON, AND V. N. VAPNIK, *A training algorithm for optimal margin classifiers*, in *Proceedings of the fifth annual workshop on Computational learning theory*, ACM, 1992, pp. 144–152.
- [7] D. M. CHANDLER, *Seven challenges in image quality assessment : past, present, and future research*, *ISRN Signal Processing*, 2013 (2013).
- [8] D. M. CHANDLER AND S. S. HEMAMI, *Vsnr : A wavelet-based visual signal-to-noise ratio for natural images*, *Image Processing, IEEE Transactions on*, 16 (2007), pp. 2284–2298.
- [9] C.-C. CHANG AND C.-J. LIN, *LIBSVM : A library for support vector machines*, *ACM Transactions on Intelligent Systems and Technology*, 2 (2011), pp. 27 :1–27 :27.
Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- [10] M. J. CHEN AND A. C. BOVIK, *No-reference image blur assessment using multiscale gradient*, in *2009 International Workshop on Quality of Multimedia Experience*, July 2009, pp. 70–74.

- [11] M.-J. CHEN, L. K. CORMACK, AND A. C. BOVIK, *No-reference quality assessment of natural stereopairs*, IEEE Transactions on Image Processing, 22 (2013), pp. 3379–3391.
- [12] M.-J. CHEN, C.-C. SU, D.-K. KWON, L. K. CORMACK, AND A. C. BOVIK, *Full-reference quality assessment of stereopairs accounting for rivalry*, Signal Processing : Image Communication, 28 (2013), pp. 1143–1155.
- [13] A. CHETOUANI, A. BEGHDAI, AND M. DERICHE, *A hybrid system for distortion classification and image quality evaluation*, Signal Processing : Image Communication, 27 (2012), pp. 948–960.
- [14] K. CHONO, Y. C. LIN, D. VARODAYAN, Y. MIYAMOTO, AND B. GIROD, *Reduced-reference image quality assessment using distributed source coding*, in 2008 IEEE International Conference on Multimedia and Expo, June 2008, pp. 609–612.
- [15] A. CIANCIO, A. L. N. T. DA COSTA, E. A. DA SILVA, A. SAID, R. SAMADANI, AND P. OBRADOR, *No-reference blur assessment of digital pictures based on multifeature classifiers*, Image Processing, IEEE Transactions on, 20 (2011), pp. 64–75.
- [16] E. COHEN AND Y. YITZHAKY, *No-reference assessment of blur and noise impacts on image quality*, Signal, Image and Video Processing, 4 (2010), pp. 289–302.
- [17] S. CORCHS, F. GASPARINI, AND R. SCETTINI, *No reference image quality classification for jpeg-distorted images*, Digital Signal Processing, 30 (2014), pp. 86–100.
- [18] J. DING AND G. SPERLING, *A gain-control theory of binocular combination*, Proceedings of the National Academy of Sciences of the United States of America, 103 (2006), pp. 1141–1146.
- [19] R. C. EBERHART AND J. KENNEDY, *A new optimizer using particle swarm theory*, in Proceedings of the sixth international symposium on micro machine and human science, vol. 1, New York, NY, 1995, pp. 39–43.
- [20] R. C. EBERHART AND Y. SHI, *Comparison between genetic algorithms and particle swarm optimization*, in Evolutionary Programming VII, Springer, 1998, pp. 611–616.
- [21] Y. FANG, K. MA, Z. WANG, W. LIN, Z. FANG, AND G. ZHAI, *No-reference quality assessment of contrast-distorted images based on natural scene statistics*, IEEE Signal Processing Letters, 22 (2015), pp. 838–842.
- [22] R. FERZLI AND L. J. KARAM, *Human visual system based no-reference objective image sharpness metric*, in Image Processing, 2006 IEEE International Conference on, IEEE, 2006, pp. 2949–2952.

-
- [23] R. FERZLI AND L. J. KARAM, *A no-reference objective image sharpness metric based on the notion of just noticeable blur (jnb)*, IEEE Transactions on Image Processing, 18 (2009), pp. 717–728.
- [24] P. GASTALDO, G. PARODI, J. REDI, AND R. ZUNINO, *No-reference quality assessment of jpeg images by using cbp neural networks*, in International Conference on Artificial Neural Networks, Springer, 2007, pp. 564–572.
- [25] S. A. GOLESTANEH AND D. M. CHANDLER, *No-reference quality assessment of jpeg images via a quality relevance map*, IEEE Signal Processing Letters, 21 (2014), pp. 155–158.
- [26] S. A. GOLESTANEH AND D. M. CHANDLER, *No-reference quality assessment of jpeg images via a quality relevance map*, Signal Processing Letters, IEEE, 21 (2014), pp. 155–158.
- [27] S. GROSSBERG AND F. KELLY, *Neural dynamics of binocular brightness perception*, Vision Research, 39 (1999), pp. 3796–3816.
- [28] L. C. H. R. SHEIKH, Z. WANG AND A. C. BOVIK, *Live image quality assessment database release 2*.
<http://live.ece.utexas.edu/research/quality>.
- [29] M. HANMANDLU, A. KUMAR, V. K. MADASU, AND P. YARLAGADDA, *Fusion of hand based biometrics using particle swarm optimization*, in Information Technology : New Generations, 2008. ITNG 2008. Fifth International Conference on, IEEE, 2008, pp. 783–788.
- [30] R. HASSAN, B. COHANIM, O. DE WECK, AND G. VENTER, *A comparison of particle swarm optimization and the genetic algorithm*, in Proceedings of the 1st AIAA multidisciplinary design optimization specialist conference, 2005, pp. 1–13.
- [31] S. S. HEMAMI AND A. R. REIBMAN, *No-reference image and video quality estimation : Applications and human-motivated design*, Signal processing : Image communication, 25 (2010), pp. 469–481.
- [32] Z. HONG, F. REN, AND Y. DING, *Saliency-based feature learning for no-reference image quality assessment*, in Green Computing and Communications (GreenCom), 2013 IEEE and Internet of Things (iThings/CPSCom), IEEE International Conference on and IEEE Cyber, Physical and Social Computing, IEEE, 2013, pp. 1790–1794.
- [33] C.-W. HSU, C.-C. CHANG, C.-J. LIN, ET AL., *A practical guide to support vector classification*, 2003.
- [34] J.-J. HWANG AND H. R. WU, *Stereo image quality assessment using visual attention and distortion predictors.*, TIIS, 5 (2011), pp. 1613–1631.

- [35] T. INSTALLATIONS AND L. LINE, *Subjective video quality assessment methods for multimedia applications*, Networks, 910 (1999), p. 37.
- [36] L. JIN, K. EGIAZARIAN, AND C.-C. J. KUO, *Perceptual image quality assessment using block-based multi-metric fusion (bmmf)*, in Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on, IEEE, 2012, pp. 1145–1148.
- [37] L. KANG, P. YE, Y. LI, AND D. DOERMANN, *Convolutional neural networks for no-reference image quality assessment*, in Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on, IEEE, 2014, pp. 1733–1740.
- [38] A. K. KEMALKAR AND V. K. BAIRAGI, *A no-reference image quality assessment*, in 2013 IEEE International Conference ON Emerging Trends in Computing, Communication and Nanotechnology (ICECCN), March 2013, pp. 462–465.
- [39] T. M. KUSUMA AND H. J. ZEPERNICK, *A reduced-reference perceptual quality metric for in-service image quality assessment*, in Mobile Future and Symposium on Trends in Communications, 2003. SympoTIC '03. Joint First Workshop on, Oct 2003, pp. 71–74.
- [40] E. C. LARSON AND D. M. CHANDLER, *Most apparent distortion : full-reference image quality assessment and the role of strategy*, Journal of Electronic Imaging, 19 (2010), p. 011006.
- [41] E. C. LARSON AND D. M. CHANDLER, *Most apparent distortion : full-reference image quality assessment and the role of strategy*, Journal of Electronic Imaging, 19 (2010), pp. 011006–011006–21.
- [42] C. LI, A. C. BOVIK, AND X. WU, *Blind image quality assessment using a general regression neural network*, IEEE Transactions on Neural Networks, 22 (2011), pp. 793–799.
- [43] F. LI, L. SHEN, D. WU, AND R. FANG, *Full-reference quality assessment of stereoscopic images using disparity-gradient-phase similarity*, in Signal and Information Processing (ChinaSIP), 2015 IEEE China Summit and International Conference on, IEEE, 2015, pp. 658–662.
- [44] Q. LI AND Z. WANG, *Reduced-reference image quality assessment using divisive normalization-based image representation*, Selected Topics in Signal Processing, IEEE Journal of, 3 (2009), pp. 202–211.
- [45] Q. LI AND Z. WANG, *Reduced-reference image quality assessment using divisive normalization-based image representation*, IEEE Journal of Selected Topics in Signal Processing, 3 (2009), pp. 202–211.

-
- [46] L. LIANG, S. WANG, J. CHEN, S. MA, D. ZHAO, AND W. GAO, *No-reference perceptual image quality metric using gradient profiles for jpeg2000*, Signal Processing : Image Communication, 25 (2010), pp. 502–516.
- [47] Y.-H. LIN AND J.-L. WU, *Quality assessment of stereoscopic 3d image compression by binocular integration behaviors*, IEEE Transactions on Image Processing, 23 (2014), pp. 1527–1542.
- [48] H. LIU, N. KLOMP, AND I. HEYNDERICKX, *A no-reference metric for perceived ringing artifacts in images*, IEEE Transactions on Circuits and Systems for Video Technology, 20 (2010), pp. 529–539.
- [49] L. LIU, B. LIU, H. HUANG, AND A. C. BOVIK, *No-reference image quality assessment based on spatial and spectral entropies*, Signal Processing : Image Communication, 29 (2014), pp. 856–863.
- [50] M. LIU AND X. YANG, *A new image quality approach based on decision fusion*, in Fuzzy Systems and Knowledge Discovery, 2008. FSKD'08. Fifth International Conference on, vol. 4, IEEE, 2008, pp. 10–14.
- [51] T.-J. LIU, W. LIN, AND C.-C. J. KUO, *Image quality assessment using multi-method fusion*, Image Processing, IEEE Transactions on, 22 (2013), pp. 1793–1807.
- [52] A. MAALOUF AND M.-C. LARABI, *Cyclop : A stereo color image quality assessment metric*, in 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2011, pp. 1161–1164.
- [53] R. A. MANAP AND L. SHAO, *Non-distortion-specific no-reference image quality assessment : A survey*, Information Sciences, 301 (2015), pp. 141–160.
- [54] P. MARZILIANO, F. DUFAUX, S. WINKLER, AND T. EBRAHIMI, *Perceptual blur and ringing metrics : application to jpeg2000*, Signal processing : Image communication, 19 (2004), pp. 163–172.
- [55] A. MCANDREW, *An introduction to digital image processing with matlab notes for scm2511 image processing*, School of Computer Science and Mathematics, Victoria University of Technology, (2004), pp. 1–264.
- [56] A. MITTAL, A. K. MOORTHY, AND A. C. BOVIK, *No-reference image quality assessment in the spatial domain*, Image Processing, IEEE Transactions on, 21 (2012), pp. 4695–4708.
- [57] A. MITTAL, R. SOUNDARARAJAN, AND A. C. BOVIK, *Making a “completely blind” image quality analyzer*, IEEE Signal Processing Letters, 20 (2013), pp. 209–212.

- [58] P. MOHAMMADI, A. EBRAHIMI-MOGHADAM, AND S. SHIRANI, *Subjective and objective quality assessment of image : A survey*, arXiv preprint arXiv :1406.7799, (2014).
- [59] A. K. MOORTHY AND A. C. BOVIK, *A two-step framework for constructing blind image quality indices*, *Signal Processing Letters, IEEE*, 17 (2010), pp. 513–516.
- [60] ———, *Blind image quality assessment : From natural scene statistics to perceptual quality*, *Image Processing, IEEE Transactions on*, 20 (2011), pp. 3350–3364.
- [61] A. K. MOORTHY, C.-C. SU, A. MITTAL, AND A. C. BOVIK, *Subjective evaluation of stereoscopic image quality*, *Signal Processing : Image Communication*, 28 (2013), pp. 870–883.
- [62] N. NARVEKAR AND L. J. KARAM, *An improved no-reference sharpness metric based on the probability of blur detection*, in *Workshop on Video Processing and Quality Metrics*, 2010.
- [63] M. NARWARIA AND W. LIN, *Svd-based quality metric for image and video using machine learning*, *Systems, Man, and Cybernetics, Part B : Cybernetics, IEEE Transactions on*, 42 (2012), pp. 347–364.
- [64] A. NINASSI, F. AUTRUSSEAU, AND P. LE CALLET, *Pseudo no reference image quality metric using perceptual data hiding*, in *Human Vision and Electronic Imaging*, 2006, pp. 146–157.
- [65] M. OSZUST, *Decision fusion for image quality assessment using an optimization approach*, *Signal Processing Letters, IEEE*, 23 (2016), pp. 65–69.
- [66] N. PONOMARENKO, O. IEREMEIEV, V. LUKIN, K. EGIАЗARIAN, L. JIN, J. ASTOLA, B. VOZEL, K. CHEHDI, M. CARLI, F. BATTISTI, ET AL., *Color image database tid2013 : Peculiarities and preliminary results*, in *Visual Information Processing (EUVIP), 2013 4th European Workshop on, IEEE*, 2013, pp. 106–111.
- [67] N. PONOMARENKO, V. LUKIN, A. ZELENSKY, K. EGIАЗARIAN, M. CARLI, AND F. BATTISTI, *Tid2008-a database for evaluation of full-reference visual quality assessment metrics*, *Advances of Modern Radioelectronics*, 10 (2009), pp. 30–45.
- [68] F. QI, D. ZHAO, AND W. GAO, *Reduced reference stereoscopic image quality assessment based on binocular perceptual information*, *IEEE Transactions on Multimedia*, 17 (2015), pp. 2338–2344.
- [69] A. REHMAN AND Z. WANG, *Reduced-reference image quality assessment by structural similarity estimation*, *Image Processing, IEEE Transactions on*, 21 (2012), pp. 3378–3389.

-
- [70] D. L. RUDERMAN AND W. BIALEK, *Statistics of natural images : Scaling in the woods*, Physical review letters, 73 (1994), pp. 814–817.
- [71] S. RYU AND K. SOHN, *No-reference quality assessment for stereoscopic images based on binocular quality perception*, IEEE Transactions on Circuits and Systems for Video Technology, 24 (2014), pp. 591–602.
- [72] M. A. SAAD, A. C. BOVIK, AND C. CHARRIER, *A dct statistics-based blind image quality index*, IEEE Signal Processing Letters, 17 (2010), pp. 583–586.
- [73] ———, *Blind image quality assessment : A natural scene statistics approach in the dct domain*, Image Processing, IEEE Transactions on, 21 (2012), pp. 3339–3352.
- [74] M. P. SAMPAT, Z. WANG, S. GUPTA, A. C. BOVIK, AND M. K. MARKEY, *Complex wavelet structural similarity : A new image similarity index*, IEEE transactions on image processing, 18 (2009), pp. 2385–2401.
- [75] D. SANTA-CRUZ, R. GROSBOIS, AND T. EBRAHIMI, *Jpeg 2000 performance evaluation and assessment*, Signal Processing : Image Communication, 17 (2002), pp. 113–130.
- [76] Z. P. SAZZAD, Y. KAWAYOKE, AND Y. HORITA, *No reference image quality assessment for jpeg2000 based on spatial features*, Signal Processing : Image Communication, 23 (2008), pp. 257–268.
- [77] Z. P. SAZZAD, S. YAMANAKA, Y. KAWAYOKEITA, AND Y. HORITA, *Stereoscopic image quality prediction*, in Quality of Multimedia Experience, 2009. QoMEX 2009. International Workshop on, IEEE, 2009, pp. 180–185.
- [78] B. SCHÖLKOPF, A. J. SMOLA, R. C. WILLIAMSON, AND P. L. BARTLETT, *New support vector algorithms*, Neural computation, 12 (2000), pp. 1207–1245.
- [79] Z. A. SEGHIR, F. HACHOUF, AND F. MORAIN-NICOLIER, *Blind image quality metric for blurry and noisy image*, in 2013 IEEE Second International Conference on Image Information Processing (ICIIP-2013), Dec 2013, pp. 193–197.
- [80] F. SHAO, S. GU, G. JANG, AND M. YU, *A novel no-reference stereoscopic image quality assessment method*, in Photonics and Optoelectronics (SOPO), 2012 Symposium on, IEEE, 2012, pp. 1–4.
- [81] F. SHAO, S. GU, G. JIANG, AND M. YU, *Stereoscopic images quality assessment by jointly evaluating image quality and depth perception*, in Fuzzy Systems and Knowledge Discovery (FSKD), 2012 9th International Conference on, IEEE, 2012, pp. 1963–1966.

- [82] F. SHAO, K. LI, W. LIN, G. JIANG, AND M. YU, *Using binocular feature combination for blind quality assessment of stereoscopic images*, Signal Processing Letters, IEEE, 22 (2015), pp. 1548–1551.
- [83] F. SHAO, K. LI, W. LIN, G. JIANG, M. YU, AND Q. DAI, *Full-reference quality assessment of stereoscopic images by learning binocular receptive field properties*, Image Processing, IEEE Transactions on, 24 (2015), pp. 2971–2983.
- [84] F. SHAO, W. LIN, S. GU, G. JIANG, AND T. SRIKANTHAN, *Perceptual full-reference quality assessment of stereoscopic images by considering binocular visual characteristics*, IEEE Transactions on Image Processing, 22 (2013), pp. 1940–1953.
- [85] F. SHAO, W. LIN, S. WANG, G. JIANG, AND M. YU, *Blind image quality assessment for stereoscopic images using binocular guided quality lookup and visual codebook*, Broadcasting, IEEE Transactions on, 61 (2015), pp. 154–165.
- [86] H. R. SHEIKH, A. C. BOVIK, AND L. CORMACK, *No-reference quality assessment using natural scene statistics : Jpeg2000*, Image Processing, IEEE Transactions on, 14 (2005), pp. 1918–1927.
- [87] H. R. SHEIKH, M. F. SABIR, AND A. C. BOVIK, *A statistical evaluation of recent full reference image quality assessment algorithms*, Image Processing, IEEE Transactions on, 15 (2006), pp. 3440–3451.
- [88] A. J. SMOLA AND B. SCHÖLKOPF, *A tutorial on support vector regression*, Statistics and computing, 14 (2004), pp. 199–222.
- [89] S. SURESH, R. V. BABU, AND H. KIM, *No-reference image quality assessment using modified extreme learning machine classifier*, Applied Soft Computing, 9 (2009), pp. 541–552.
- [90] H. TANG, N. JOSHI, AND A. KAPOOR, *Learning a blind measure of perceptual image quality*, in Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on, IEEE, 2011, pp. 305–312.
- [91] ———, *Blind image quality assessment using semi-supervised rectifier networks*, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 2877–2884.
- [92] Z. TAOJIA, Z. HONG, AND Y. DING, *A no-reference quality metric for blur image*, in Green Computing and Communications (GreenCom), 2013 IEEE and Internet of Things (iThings/CPSCom), IEEE International Conference on and IEEE Cyber, Physical and Social Computing, IEEE, 2013, pp. 1813–1816.

- [93] Z. TAOJIA, Z. HONG, AND Y. DING, *A no-reference quality metric for blur image*, in 2013 IEEE International Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber, Physical and Social Computing, Aug 2013, pp. 1813–1816.
- [94] P. V. VU AND D. M. CHANDLER, *A no-reference quality assessment algorithm for jpeg2000-compressed images based on local sharpness*, in IS&T/SPIE Electronic Imaging, International Society for Optics and Photonics, 2013, pp. 865302–865302.
- [95] J. WANG, A. REHMAN, K. ZENG, S. WANG, AND Z. WANG, *Quality prediction of asymmetrically distorted stereoscopic 3d images*, IEEE Transactions on Image Processing, 24 (2015), pp. 3400–3414.
- [96] J. WANG, K. ZENG, AND Z. WANG, *Quality prediction of asymmetrically distorted stereoscopic images from single views*, in 2014 IEEE International Conference on Multimedia and Expo (ICME), IEEE, 2014, pp. 1–6.
- [97] Z. WANG, *Applications of objective image quality assessment methods [applications corner]*, IEEE Signal Processing Magazine, 28 (2011), pp. 137–142.
- [98] Z. WANG AND A. C. BOVIK, *A universal image quality index*, IEEE Signal Processing Letters, 9 (2002), pp. 81–84.
- [99] Z. WANG AND A. C. BOVIK, *Modern image quality assessment*, Synthesis Lectures on Image, Video, and Multimedia Processing, 2 (2006), pp. 1–156.
- [100] Z. WANG, A. C. BOVIK, AND B. L. EVAN, *Blind measurement of blocking artifacts in images*, in Proceedings 2000 International Conference on Image Processing (Cat. No.00CH37101), vol. 3, 2000, pp. 981–984 vol.3.
- [101] Z. WANG, A. C. BOVIK, H. R. SHEIKH, AND E. P. SIMONCELLI, *Image quality assessment : from error visibility to structural similarity*, Image Processing, IEEE Transactions on, 13 (2004), pp. 600–612.
- [102] Z. WANG, H. R. SHEIKH, AND A. C. BOVIK, *No-reference perceptual quality assessment of jpeg compressed images*, in Image Processing. 2002. Proceedings. 2002 International Conference on, vol. 1, IEEE, 2002, pp. I–477.
- [103] Z. WANG, E. P. SIMONCELLI, AND A. C. BOVIK, *Multiscale structural similarity for image quality assessment*, in Signals, Systems and Computers, 2004. Conference Record of the Thirty-Seventh Asilomar Conference on, vol. 2, IEEE, 2003, pp. 1398–1402.
- [104] H. R. WILSON, *Quantitative characterization of two types of line-spread function near the fovea*, Vision Research, 18 (1978), pp. 971–981.

- [105] H. R. WILSON AND J. R. BERGEN, *A four mechanism model for threshold spatial vision*, Vision research, 19 (1979), pp. 19–32.
- [106] H. R. WU AND M. YUEN, *A generalized block-edge impairment metric for video coding*, IEEE Signal Processing Letters, 4 (1997), pp. 317–320.
- [107] Q. XU, G. ZHAI, M. LIU, AND K. GU, *Using structural degradation and parallax for reduced-reference quality assessment of 3d images*, in 2014 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting, IEEE, 2014, pp. 1–6.
- [108] G. YAMMINE, E. WIGE, AND A. KAUP, *A no-reference blocking artifacts visibility estimator in images*, in 2010 IEEE International Conference on Image Processing, Sept 2010, pp. 2497–2500.
- [109] Q. YAN, Y. XU, AND X. YANG, *No-reference image blur assessment based on gradient profile sharpness*, in 2013 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB), June 2013, pp. 1–4.
- [110] J. YANG, C. HOU, Y. ZHOU, Z. ZHANG, AND J. GUO, *Objective quality assessment method of stereo images*, in 2009 3DTV Conference : The True Vision-Capture, Transmission and Display of 3D Video, IEEE, 2009, pp. 1–4.
- [111] P. YE AND D. DOERMANN, *No-reference image quality assessment using visual codebooks*, IEEE Transactions on Image Processing, 21 (2012), pp. 3129–3138.
- [112] P. YE, J. KUMAR, L. KANG, AND D. DOERMANN, *Unsupervised feature learning framework for no-reference image quality assessment*, in Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on, IEEE, 2012, pp. 1098–1105.
- [113] J. YOU, L. XING, A. PERKIS, AND X. WANG, *Perceptual quality assessment for stereoscopic images based on 2d image quality metrics and disparity analysis*, in Proc. of International Workshop on Video Processing and Quality Metrics for Consumer Electronics, Scottsdale, AZ, USA, 2010.
- [114] J. ZHANG AND T. M. LE, *A new no-reference quality metric for jpeg2000 images*, Consumer Electronics, IEEE Transactions on, 56 (2010), pp. 743–750.
- [115] J. ZHANG, S. H. ONG, AND T. M. LE, *Kurtosis-based no-reference quality assessment of jpeg2000 images*, Signal Processing : Image Communication, 26 (2011), pp. 13–23.
- [116] L. ZHANG, L. ZHANG, X. MOU, AND D. ZHANG, *Fsim : A feature similarity index for image quality assessment*, IEEE transactions on Image Processing, 20 (2011), pp. 2378–2386.
- [117] X. ZHANG, X. FENG, W. WANG, AND W. XUE, *Edge strength similarity for image quality assessment*, IEEE Signal processing letters, 20 (2013), pp. 319–322.

- [118] Y. ZHANG AND D. M. CHANDLER, *No-reference image quality assessment based on log-derivative statistics of natural scenes*, *Journal of Electronic Imaging*, 22 (2013), pp. 043025–043025.
- [119] Y. ZHANG, A. K. MOORTHY, D. M. CHANDLER, AND A. C. BOVIK, *C-diivine : No-reference image quality assessment based on local magnitude and phase statistics of natural scenes*, *Signal Processing : Image Communication*, 29 (2014), pp. 725–747.
- [120] W. ZHOU, G. JIANG, M. YU, F. SHAO, AND Z. PENG, *Pmfs : a perceptual modulated feature similarity metric for stereoscopic image quality assessment*, *IEEE Signal Processing Letters*, 21 (2014), pp. 1003–1006.